**Title**
Designing Rational Combination Strategies for Overcoming Drug Resistance in Breast Cancer



**Author**
Donnella, Hayley Jeton

**Publication Date**
2019

Designing Rational Combination Strategies for Overcoming Drug Resistance in Breast Cancer

by
Hayley Donnella

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Pharmaceutical Sciences and Pharmacogenomics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Sourav Bandyopadhyay
F1E602E27884493...
Sourav Bandyopadhyay
Chair

DocuSigned by:
Kevan Shokat
Kevan Shokat

DocuSigned by:
Andrei Goga
2582F8BD11124B6...
Andrei Goga

Committee Members

# Dedication

For J.L.N. and to all those enduring the battle.

# Acknowledgements

I could write a thesis in and of itself attempting to express my gratitude for all those involved who have made the completion of this dissertation a reality. It truly takes a village, and a special kind of crazy, to endure the path to a PhD – and while no magnitude of words can truly capture how grateful I am, I will try to keep this short and sweet:

I would like to thank my advisor, Sourav Bandyopadhyay, for his mentorship and guidance along this journey. Sourav's continued support, knowledge, and scientific curiosity has been a strong source of encouragement during this pursuit. I am forever grateful for the freedom I was given in my research, allowing me to develop projects independently and foster ideas on my own accord. I would not be the scientist I am today without his leadership, and could not have envisioned a better training environment for this process.

I would like to thank other scientific mentors and colleagues I've relied on during this time. Andrei Goga, Kevan Shokat, Deanna Kroetz, John Gordan, Trever Bivona, Allan Balmain, Rob Judson, Saroja Ramanujan, Joya Chandra, Melissa Singh, members of the Bandyopadhyay lab past and present (particularly James Webber who entertained my bouts of panic and frazzlement as fresh grad student), Emmalyn Chen, as well as my entire PSPG cohort have shaped my thinking and experiences and provided a strong foundation to build my scientific career from.

I would like to thank my friends and family. I will never be able to fully convey the magnitude of love and support I've received from my parents, Paula and Jay Donnella. I will always appreciate how they've given me the confidence and every opportunity needed to succeed and been an endless source of support in every facet of my life. There simply are no words here. This extends to my entire family including my sisters Heather and Jayla, my grandparents, my aunt Nikki, and my newly-extended family of in-laws who have all played a substantial role in entertaining my endeavors and believed in me over the course of this graduate work. Thank you to all of you.

I would like to thank my husband, Ryan Woods. This truly would not have been possible without his incredibly fierce love and support during this time and beyond. I am exceedingly lucky to have found a partner with an equally matched determination in life. From our infamous weekend work dates, to being my voice of reason and confidence when my patience was wearing thin – Ryan has earned his own PhD twice over for the work he has done as he Patiently helped Donnella [get through her] PhD. And with that, Ryan Woods $PhD^2$, I love you.

# Contributions

Several chapters in this dissertation contain material that has been published previously. They do not necessarily represent the final published form and in most cases have been edited slightly.

Chapter 2 of this thesis is a reprint of previously published work as it was accepted in the scientific journal *Nature Chemical Biology* in June 2018:

HJ Donnella, JT Webber, RS Levin, R Camarda, O Momcilovic, N Bayani, KN Shah, JE Korkola, KM Shokat, A Goga, JD Gordan, S Bandyopadhyay. "Kinome rewiring reveals AURKA limits PI3K-pathway inhibitor efficacy in breast cancer." Nature Chemical Biology, 2018. doi:10.1038/s41589-018-0081-9

H.J.D., J.T.W., K.M.S, J.K., J.D.G., and S.B. contributed towards study conceptualization. H.J.D., J.T.W., and J.D.G. performed data analyses supporting the study. H.J.D. designed and performed the majority of experiments. N.B. assisted with samples for initial MIBs/MS profiling. R.S.L. J.T.W., and J.D.G. provided technical advice and guided the interpretation of mass spectrometry data from MIB/MS profiling. R.C. and O.M. assisted with animal studies, and K.N.S. helped with additional experiments. H.J.D. and S.B. composed the original draft and all authors contributed towards manuscript finalization. S.B. and A.G. supervised the study.

Chapter 3 of this thesis is a partial reprint of material currently in submission for publication in a scientific journal and has been uploaded to bioRxiv:

HJ Donnella, S Kaushik, A Tambe, JT Webber, O Momcilovic, A Goga, S Bandyopadhyay. "Single-cell pharmacogenomics targets an inflamed chemoresistant tumor subpopulation in triple negative breast cancer."

H.J.D. and S.B. contributed towards study conceptualization. H.J.D., A.T., and S.K. performed data analyses supporting the study. H.J.D. designed and performed the majority of experiments. O.M. assisted with samples for initial single-cell sequencing. H.J.D. and S.B. composed the original draft and all authors contributed towards manuscript finalization. S.B. and A.G. supervised the study.

# Designing Rational Combination Strategies for Overcoming Drug Resistance in Breast Cancer

Hayley Jeton Donnella

Abstract

Drug resistance is a ubiquitous problem in the therapeutic management of breast cancer, even in the context of next-generation targeted therapies where only modest clinical improvements have been observed despite a tumors mutational load for a given target pathway or intrinsic subtype. To devise effective anti-cancer treatment strategies, new systems-based methods are needed to fully interpret factors underlying drug responses encompassing both genetic and non-genetic mechanisms. Here we developed two approaches towards designing novel combination strategies for overcoming drug resistance. First, using an unbiased chemoproteomics approach, we profiled kinome dynamics across breast cancer cells in response to various targeted therapies and identified signaling changes that correlate with drug sensitivity. This signaling map identified survival factors whose presence limits the efficacy of targeted therapies and revealed AURKA as a new co-targeting opportunity to enhance the therapeutic efficacy of PI3K-pathway inhibitors in breast cancer. Second, we used single-cell transcriptomics data and pharmacogenomic modeling as a way to inform upfront drug combinations based on systematic analysis of tumor subpopulation architectures. Using *in silico* and experimental approaches, our study provides an effective new framework to discover drug combinations capable of counteracting intrinsic cell

variability by predicting drug responses of single cells within tumor cell subpopulations and systematically links transcriptional heterogeneity with drug actionability to optimize therapy combinations.

# Table of Contents

# List of Figures

# List of Datasets

**Supplementary Dataset 2.1** Kinase activity scores and *P* values from MIBs/MS profiling.

**Supplementary Dataset 2.2** Gene set enrichment analysis (GSEA) of BYL719-treated cells.

**Supplementary Dataset 2.3** Receptor status, synergy scores (Loewe and Bliss) and assessment of increase in apoptosis over single-agent or combinations tested in MIBs/MS study.

**Supplementary Dataset 2.4** MYC signature derived from isogenic MCF10A cell lines.

**Supplementary Dataset 3.1** Single-cell RNA sequencing metrics summary.

**Supplementary Dataset 3.2** Differentially expressed genes by cluster for each breast cancer cell line and patient-derived xenograft model sequenced.

**Supplementary Dataset 3.3** Gene set enrichment analysis (GSEA) of $MCF7^{ER+}$ Cluster 7 and non-Cluster 7 cells.

**Supplementary Dataset 3.4** Proliferation, ESR1/ER, ERBB2/HER2, EMT, Basal, and Interferon-stimulated gene (ISG) module gene sets.

# Chapter 1.

# Introduction

## 1.1 Breast cancer landscape and treatment options

Breast cancer remains a highly prevalent disease with 1 in 8 women developing invasive breast cancer during her lifetime. This presents as the most common cancer among women in the United States, accounting for over 15% of all newly diagnosed cancers, with a predicted 258,600 new cases and 41,670 breast cancer patient deaths in 2019[1]. Breast cancers are typically diagnosed based on the presence/absence of receptor expression into receptor-positive (estrogen receptor, ER; progesterone receptor, PR), ErbB2 (HER2)-amplified or triple-negative breast cancer (TNBC) indications, and treatment options vary depending on disease stage at presentation and tumor subtype. However, as a result of inter-tumoral heterogeneity (between patient variability in tumor phenotype), improvements in the clinical management of this disease have not been equal across all patient groups. In breast cancer, the discovery of biomarkers has now led to classifying tumors based on gene expression patterns into five molecular subtypes including Luminal A/B, HER2-enriched, and two triple-negative subtypes, Basal-like and Claudin-low[2–4]. Molecular variability between individual cancer patients can predict who will respond to treatment, such as differences in molecular subtypes, which is crucial in determining predicted responsiveness to standard-of-care therapies[5–7]. Accordingly, classification of intrinsic molecular subtype from bulk patient tumor samples has become an important diagnostic tool to

predict risk of relapse and response to chemotherapy in breast cancers. However, patients do not present with homogenous tumors clinically, often times lacking uniform expression of therapeutic targets[8]. This heterogeneity can severely impact the clinical assessment of a tumor and subsequent therapeutic regimens, and may facilitate recurrence whereby subpopulations of cells are spared during treatment. Furthermore, several preclinical and clinical observations have shown that genotype alone fails to predict drug response[9–12], suggesting that therapeutic efficacy is often based on the presence of molecular modifiers not represented by global measures of subtype and receptor diversity.

## 1.2 Therapeutic targeting of driver kinases in breast cancer

As our understanding of the molecular biology of cancer has advanced, drug development has shifted towards agents that target specific molecular alterations in tumors. Exome sequencing has revealed that frequent mutation in genes of the PI3K/AKT/mTOR pathway is a prominent feature of breast cancers. Genomic alterations of PI3K-pathway components, including *PTEN*, *PIK3CA*, and *AKT1*, occur in over 60% of breast malignancies[13]. Accordingly, inhibition of this signaling pathway is an attractive avenue of investigation for therapeutic intervention and has led to the advancement of more than 30 inhibitors currently in clinical development that target the PI3K/AKT/mTOR machinery in breast and other cancers[14–18]. Despite this being one of the most active areas of drug development, multiple clinical trials demonstrate that monotherapy responses to these inhibitors remain modest. Collectively, there was merely an 18% response rate

in patients harboring *PIK3CA* mutations or PTEN aberrations and a minimal 1.8 month improvement of progression free survival (PFS) when treated with single-agent PI3K/AKT/mTOR inhibitors[19]. It is therefore likely that new approaches to identifying efficacious treatment regimens are necessary in order to improve clinical outcomes.

## 1.3 Factors underlying drug resistance and response to targeted therapies in breast cancer

Drug resistance is a major problem in breast cancer where insensitivity to chemotherapy is associated with a 40 to 80% risk of recurrence after neo-adjuvant therapy resulting in incurable metastatic disease and death for most patients[20–22]. The need for new therapeutic approaches is particularly pressing in the context of triple-negative breast cancers (TNBC) where responses to chemotherapy are often not durable. In the case of neo-adjuvant chemotherapy, 30 to 50% of patients with localized TNBC evolve resistance leading to poor overall survival for this group of patients[23,24]. The ubiquity of drug resistance in breast cancer remains a problem even in the context of next-generation targeted therapies where only modest improvements in PFS and overall survival (OS) have been observed despite a tumors mutational load for a given target pathway or intrinsic subtype[11,12,19,25,26]. As such, new approaches are needed to fully interpret factors underlying drug responses in breast cancer encompassing both genetic and non-genetic mechanisms.

### 1.3.1 The kinome resiliency conundrum and adaptations to targeted therapies

The clinical observation that tumors with mutations in the PI3K/AKT/mTOR pathway fail to respond to therapy suggests that additional factors modulate cellular response and drive resistance to PI3K-pathway inhibition. Such factors include (1) drug-induced changes in feedback signaling resulting in pathway reactivation, and (2) compensatory activation of parallel signaling networks capable of transmitting new oncogenic signals (Fig. 1.1). Within the PI3K/AKT/mTOR signaling network, numerous drug-induced feedback adaptations are now defined. For example, treatment of PI3K-hyperactivated cells with an allosteric mTORC1 inhibitor induces upstream receptor tyrosine kinase (RTK) signaling resulting in increased AKT activation[27,28]. mTORC1 inhibition is also known to directly induce PI3K signaling by relieving GRB10-mediated suppression[29,30]. Similarly, AKT inhibition (and direct inhibitors of PI3K to the extent that they result in AKT inhibition) induces the RNA expression of RTKs, thus relieving feedback suppression of RTK activity[31,32]. These data indicate drug-induced reliefs of feedback result in reactivation of the PI3K/AKT/mTOR pathway, effectively rendering drugs directed towards this pathway less potent. An alternate means of targeted therapy resistance is drug-induced compensatory activation of alternate signaling pathways (2). The PI3K and RAS/RAF/MEK/ERK networks interact at multiple levels; several groups have shown inhibition of PI3K signaling nodes results in MAPK pathway activation and vice versa[31,33–37]. These studies collectively highlight the emergence of drug-induced resistance and underscore the importance of redundancy and feedback within signaling networks.

Discovery of factors underlying drug-induced feedback phenomena has guided the development of current combination therapies focused on pursuing specific downstream effectors of the target pathway, or components of compensatory pathways as companion targets. In the context of breast cancer, multiple efforts in the field have identified mTORC1 as a survival factor whose suppression is necessary for PI3K-pwaty inhibitor sensitivity[38,39]. This observation has led to multiple clinical trials directed towards using drug combinations with PI3K/AKT/mTOR inhibitors, yet reported results have yielded suboptimal outcomes due to increased systemic toxicity and cytostatic tumor effects[40]. Hence, there remains a pressing need to uncover routes of resistance and combination targets that have not yet been explored in order to improve therapeutic efficiency. As such, identifying additional survival factors requires a comprehensive understanding of signaling dynamics in response to treatment and insight as to how these dynamics contribute to drug resistance.

### 1.3.2 Profiling kinome activation dynamics using Multiplexed Inhibitor Beads (MIBs)

To do so, we need an unbiased approach to monitor dynamic signaling responses following targeted kinase inhibition. Initial efforts in kinome adaptation studies made use of kinase activity assays, immunoblotting for select kinases and their active forms, or large-scale phosphoproteomic analyses[41,42]. However, these studies are limited due to reagent availability and specificity, and the complexity of the phosphoproteome without preference for kinases. To address this, kinase enrichment strategies using chemoproteomic techniques that combine kinase affinity capture with quantitative mass spectrometry (MS) have emerged. Investigators used

immobilized Type I kinase inhibitors coupled to Sepharose beads to target the ATP-binding pocket[43–45] (Fig. 1.2a). This allows for unbiased capture of endogenous kinases based on activity state and abundance. Recently, a derivation of this approach showed that capture of expressed kinases can be improved by multiplexing a set of immobilized pan-kinase inhibitor beads (MIBs) onto a modular affinity column[46–48]. MIB-bound kinases can then be identified by off-line LC separation and quantitative MS, providing a systems-wide approach to profile kinome activity (Fig. 1.2b,c).

### 1.3.3 Rare tumor subpopulations and heterogeneity at the single-cell transcript level

Another source of drug resistance in breast cancer can be attributed to the presence of intra-tumoral heterogeneity and rare subpopulations of cells having differential impacts on overall response (Fig. 1.3). Pre-existing genetically diverse subclones within a tumor mass can harbor resistance-conferring mutations resulting in eventual clonal expansion following a Darwinian selection of distinct cell subpopulations after an initial response to therapy. However, clinical evidence from breast cancers undergoing neo-adjuvant chemotherapy have indicated treatment often does not result in subclonal selection demonstrating that it is not necessarily genetically encoded[49–52]. In the absence of pre-existing subclones, subpopulations of cells can undergo a drug-induced adaptation towards a drug-tolerant state where cells can persist over the course of treatment and eventually acquire a variety of possible resistance mechanisms including secondary mutations. Alternatively, rare subpopulations of cells in different cellular states can be pre-existing prior to therapy and harbor differential drug sensitivities leading to distinct

subpopulations achieving therapeutic escape. Hence new approaches are needed to counteract both genetic and non-genetic factors contributing to drug resistance in breast cancer.

### 1.3.4 Single-cell approaches for characterizing heterogeneity in breast cancer

Single-cell DNA- and RNA-sequencing methods have emerged as powerful tools for resolving intra-tumoral heterogeneity, reconstructing tumor evolutionary lineages, and characterizing distinct tumor cell subpopulations in breast cancers[49,53,54]. However, approaches to use these data derived from single-cell studies towards predicting and testing new therapeutic strategies have remained a challenge. Studies linking genomics to drug efficacy, or pharmacogenomics, have relied on the analysis of large panels of cell lines or PDX systems having been both molecularly profiled and functionally challenged through high-throughput drug screens. However, these bulk assays have limited application to studying the functional consequences of cell-cell variability in that heterogeneous responses of cells from within a bulk population and subtle changes in cell state or gene expression are obscured in bulk analysis. Therefore, the functional implications that individual cells within a population pose on therapeutic responses remains unknown due to the lack in feasibility of directly challenging a single live cell with drug. As such, approaches for characterizing genetic and functional variability contributing to drug response at the cellular level are needed to inform upfront drug combinations that elicit more complete anti-tumor responses.

## 1.4 Overarching project goals

To date, there remain no gold-standard methods for designing rational combination strategies capable of combating drug resistance in breast cancer. This body of work seeks to address this challenge in a two-pronged approach studying various non-genetic factors contributing to drug response.

(1) As detailed in Chapter 2, our first project centered on the concept of kinome signaling and pathway dynamics altering drug response. The guiding hypothesis of this research is that intrinsic resistance mechanisms for PI3K-directed therapeutics might coalesce on a distinct set of critical effector nodes. Mapping patterns of kinome behavior globally provides valuable, unbiased insights into the mechanistic implications of cell cycle and response to PI3K-pathway targeted agents. This method of analysis may provide a more precise model of cell signaling responses, enabling the accurate prediction of viable targets and thus constitutes a novel, translatable approach to developing rational therapeutic combinations for breast cancer treatment. In this study we developed a kinome-profiling technique to map signaling changes in response to various target therapies across subtypes and identified AURKA as a new co-targeting opportunity to enhance the therapeutic efficacy of PI3K-pathway inhibitors in breast cancer.

(2) As detailed in Chapter 3, the second project in this body of work involved delineating the impact of transcriptional tumor cell heterogeneity on treatment response. We

hypothesized that non-genetic differences in cellular states between individual cancer cells may also reflect altered drug sensitivities and as such, can be used to inform upfront drug combinations that elicit more complete anti-tumor responses. Through pharmacogenomic modeling based on bulk cancer cell line drug sensitivities, we developed a new approach using a Modular Drug Response Prediction Pipeline ($MOD_{Rx}$) in order to systematically discover drug combinations capable of counteracting tumor heterogeneity by predicting drug responses of individual breast cancer cells. Using *in silico* and experimental approaches, our study provides an effective new framework to discover drug combinations capable of counteracting intrinsic cell variability by predicting drug responses of tumor cell subpopulations and systematically links transcriptional heterogeneity with drug actionability to optimize therapy combinations.

## 1.5 Figures

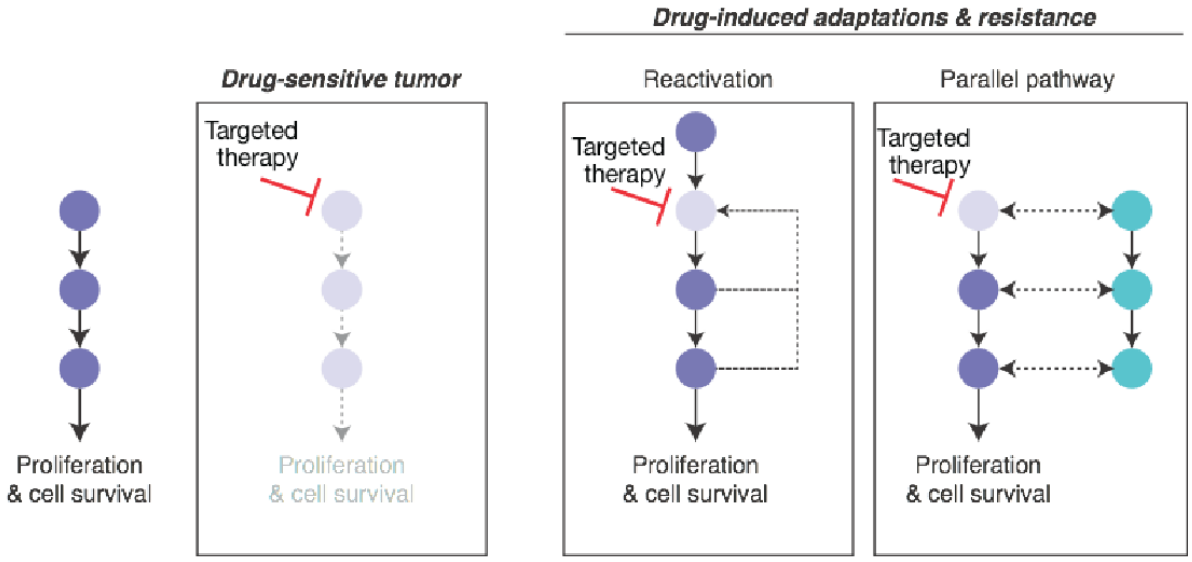



**Figure 1.1 Kinome resiliency and adaptations mechanisms.** Schematic detailing signaling dynamics in response to treatment with targeted therapies for a drug-sensitive tumor (left) and drug-resistant tumor setting (right). Signaling network crosstalk and redundancies can lead to drug-induced adaptations including reactivation of the target pathway and compensatory activation of alternate, parallel pathways.

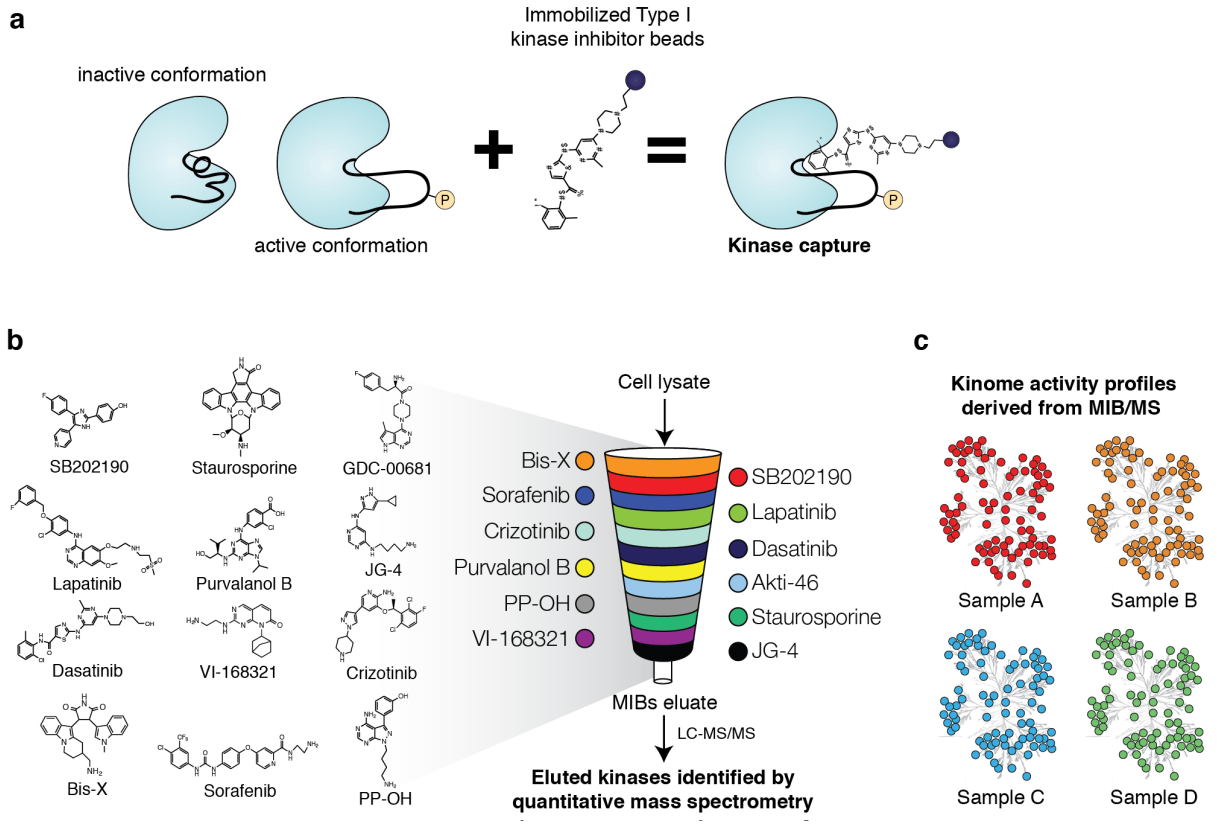



**Figure 1.2 Measuring kinome activation dynamics using Multiplexed Inhibitor Beads (MIBs) technology.** (**a**) Type I kinase inhibitors (ATP-analogs) immobilized onto beads can be used to immunoprecipitate active kinases through binding of their ATP-binding pocket providing an unbiased way to capture endogenous kinases based on activity state and abundance. (**b**) Multiplexing a panel of immobilized-inhibitor beads onto an affinity column allows for broad coverage of the kinome. Eluted kinases are then identified by LC-MS/MS to generate kinome activity profiles. (**c**) MIBs/MS can be used to generate activity profiles across several different cell types or sample conditions providing a useful tool to characterize kinome dynamics.

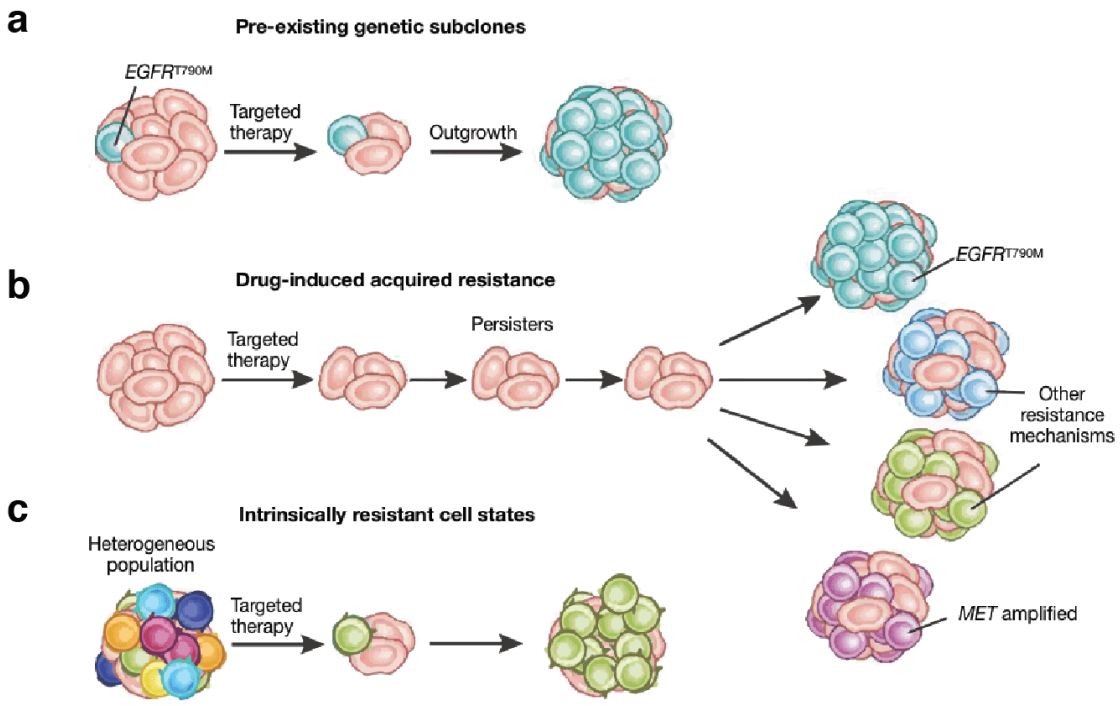



**Figure 1.3 Rare tumor subpopulations and heterogeneity at the single-cell transcript level.** Figure adapted from Oxnard, GR.[55] doi: 10.1038/nm.4058. (**a**) Pre-existing genetic subclones within a tumor mass can harbor resistance-conferring mutations resulting in eventual clonal expansion after an initial response to targeted therapy. (**b**) Targeted therapy treatment induces a durable drug-tolerant state where cells can persist over the course of treatment and eventually acquire a variety of possible resistance mechanisms including secondary mutations. (**c**) Alternatively, rare subpopulations of cells in different cellular states can be pre-existing prior to therapy and harbor differential drug sensitivities leading to distinct subpopulations achieving therapeutic escape.

## 1.6 References


1. National Cancer Institute, DCCPS, Surveillance Research Program. Surveillance, Epidemiology, and End Results (SEER) Program Populations (1969-2017). Cancer of the Female Breast. *SEER* www.seer.cancer.gov.
2. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
3. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).
4. Lehmann, B. D. *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* **121**, 2750–2767 (2011).
5. Ayers, M. *et al.* Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **22**, 2284–2293 (2004).
6. Sørlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10869–10874 (2001).
7. Parker, J. S. *et al.* Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).
8. Rivenbark, A. G., O'Connor, S. M. & Coleman, W. B. Molecular and cellular heterogeneity in breast cancer: challenges for personalized medicine. *Am. J. Pathol.* **183**, 1113–1124 (2013).
9. Slamon, D. J. *et al.* Use of Chemotherapy plus a Monoclonal Antibody against HER2 for Metastatic Breast Cancer That Overexpresses HER2. *N. Engl. J. Med.* **344**, 783–792 (2001).

10. Turner, N. C. *et al.* Palbociclib in Hormone-Receptor–Positive Advanced Breast Cancer. *N. Engl. J. Med.* **373**, 209–219 (2015).

11. Daemen, A. *et al.* Modeling precision treatment of breast cancer. *Genome Biol.* **14**, R110 (2013).

12. Heiser, L. M. *et al.* Subtype and pathway specific responses to anticancer compounds in breast cancer. *Proc. Natl. Acad. Sci.* **109**, 2724–2729 (2012).

13. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).

14. Millis SZ, Ikeda S, Reddy S, Gatalica Z & Kurzrock R. Landscape of phosphatidylinositol-3-kinase pathway alterations across 19 784 diverse solid tumors. *JAMA Oncol.* (2016) doi:10.1001/jamaoncol.2016.0891.

15. Fleuren, E. D. G., Zhang, L., Wu, J. & Daly, R. J. The kinome 'at large' in cancer. *Nat. Rev. Cancer* **16**, 83–98 (2016).

16. Dey, N., Leyland-Jones, B. & De, P. MYC-xing it up with PIK3CA mutation and resistance to PI3K inhibitors: summit of two giants in breast cancers. *Am. J. Cancer Res.* **5**, 1–19 (2014).

17. Rodon, J., Dienstmann, R., Serra, V. & Tabernero, J. Development of PI3K inhibitors: lessons learned from early clinical trials. *Nat. Rev. Clin. Oncol.* **10**, 143–153 (2013).

18. Zardavas, D., Baselga, J. & Piccart, M. Emerging targeted agents in metastatic breast cancer. *Nat. Rev. Clin. Oncol.* **10**, 191–210 (2013).

19. Janku, F. *et al.* Assessing PIK3CA and PTEN in Early-Phase Trials with PI3K/AKT/mTOR Inhibitors. *Cell Rep.* **6**, 377–387 (2014).

20. Cortazar, P. *et al.* Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet Lond. Engl.* **384**, 164–172 (2014).

21. von Minckwitz, G. *et al.* Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **30**, 1796–1804 (2012).

22. Symmans, W. F. *et al.* Long-Term Prognostic Risk After Neoadjuvant Chemotherapy Associated With Residual Cancer Burden and Breast Cancer Subtype. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **35**, 1049–1060 (2017).

23. Foulkes, W. D., Smith, I. E. & Reis-Filho, J. S. Triple-negative breast cancer. *N. Engl. J. Med.* **363**, 1938–1948 (2010).

24. Liedtke, C. *et al.* Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **26**, 1275–1281 (2008).

25. Untch, M. *et al.* Neoadjuvant treatment with trastuzumab in HER2-positive breast cancer: results from the GeparQuattro study. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **28**, 2024–2031 (2010).

26. Pierga, J.-Y. *et al.* A multicenter randomized phase II study of sequential epirubicin/cyclophosphamide followed by docetaxel with or without celecoxib or trastuzumab according to HER2 status, as primary chemotherapy for localized invasive breast cancer patients. *Breast Cancer Res. Treat.* **122**, 429–437 (2010).

27. O'Reilly, K. E. *et al.* mTOR Inhibition Induces Upstream Receptor Tyrosine Kinase Signaling and Activates Akt. *Cancer Res.* **66**, 1500–1508 (2006).

28. Shi, Y., Yan, H., Frost, P., Gera, J. & Lichtenstein, A. Mammalian target of rapamycin inhibitors activate the AKT kinase in multiple myeloma cells by up-regulating the insulin-like growth factor receptor/insulin receptor substrate-1/phosphatidylinositol 3-kinase cascade. *Am. Assoc. Cancer Res.* **4**, 1533–1540 (2005).

29. Hsu, P. P. *et al.* The mTOR-Regulated Phosphoproteome Reveals a Mechanism of mTORC1-Mediated Inhibition of Growth Factor Signaling. *Science* **332**, 1317–1322 (2011).

30. Yu, Y. *et al.* Phosphoproteomic Analysis Identifies Grb10 as an mTORC1 Substrate That Negatively Regulates Insulin Signaling. *Science* **332**, 1322–1326 (2011).

31. Chandarlapaty, S. *et al.* AKT Inhibition Relieves Feedback Suppression of Receptor Tyrosine Kinase Expression and Activity. *Cancer Cell* **19**, 58–71 (2011).

32. Chakrabarty, A., Sánchez, V., Kuba, M. G., Rinehart, C. & Arteaga, C. L. Feedback upregulation of HER3 (ErbB3) expression and activity attenuates antitumor effect of PI3K inhibitors. *Proc. Natl. Acad. Sci.* **109**, 2718–2723 (2012).

33. Mendoza, M. C., Er, E. E. & Blenis, J. The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. *Trends Biochem. Sci.* **36**, 320–328 (2011).

34. Mirzoeva, O. K. *et al.* Basal Subtype and MAPK/ERK Kinase (MEK)-Phosphoinositide 3-Kinase Feedback Signaling Determine Susceptibility of Breast Cancer Cells to MEK Inhibition. *Cancer Res.* **69**, 565–572 (2009).

35. Carracedo, A. *et al.* Inhibition of mTORC1 leads to MAPK pathway activation through a PI3K-dependent feedback loop in human cancer. *J. Clin. Invest.* (2008) doi:10.1172/JCI34739.

36. Rodrik-Outmezguine, V. S. *et al.* Overcoming mTOR resistance mutations with a new-generation mTOR inhibitor. *Nature* **534**, 272–276 (2016).

37. Serra, V. *et al.* PI3K inhibition results in enhanced HER signaling and acquired ERK dependency in HER2-overexpressing breast cancer. *Oncogene* **30**, 2547–2557 (2011).

38. Elkabets, M. *et al.* mTORC1 Inhibition Is Required for Sensitivity to PI3K p110α Inhibitors in PIK3CA-Mutant Breast Cancer. *Sci. Transl. Med.* **5**, 196ra99-196ra99 (2013).

39. Vora, S. R. *et al.* CDK 4/6 Inhibitors Sensitize PIK3CA Mutant Breast Cancer to PI3K Inhibitors. *Cancer Cell* **26**, 136–149 (2014).

40. Shah, P. D. & Chandarlapaty, S. Resistance to PI3K Pathway Inhibition. in *PI3K-mTOR in Cancer and Cancer Therapy* (eds. Dey, N., De, P. & Leyland-Jones, B.) 125–147 (Springer International Publishing, 2016). doi:10.1007/978-3-319-34211-5_5.

41. Olsen, J. V. *et al.* Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* **127**, 635–648 (2006).

42. Anastassiadis, T., Deacon, S. W., Devarajan, K., Ma, H. & Peterson, J. R. Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity. *Nat. Biotechnol.* **29**, 1039–1045 (2011).

43. Godl, K. *et al.* An efficient proteomics method to identify the cellular targets of protein kinase inhibitors. *Proc. Natl. Acad. Sci.* **100**, 15434–15439 (2003).

44. Bantscheff, M. *et al.* Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat. Biotechnol.* **25**, 1035–1044 (2007).

45. Daub, H. Quantitative Proteomics of Kinase Inhibitor Targets and Mechanisms. *ACS Chem. Biol.* **10**, 201–212 (2015).

46. Duncan, J. S. *et al.* Dynamic Reprogramming of the Kinome in Response to Targeted MEK Inhibition in Triple-Negative Breast Cancer. *Cell* **149**, 307–321 (2012).

47. Cooper, M. J. *et al.* Application of Multiplexed Kinase Inhibitor Beads to Study Kinome Adaptations in Drug-Resistant Leukemia. *PLoS ONE* **8**, e66755 (2013).

48. Stuhlmiller, T. J. *et al.* Inhibition of Lapatinib-Induced Kinome Reprogramming in ERBB2-Positive Breast Cancer by Targeting BET Family Bromodomains. *Cell Rep.* **11**, 390–404 (2015).

49. Kim, C. *et al.* Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing. *Cell* **173**, 879-893.e13 (2018).

50. Yates, L. R. *et al.* Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat. Med.* **21**, 751–759 (2015).

51. Balko, J. M. *et al.* Molecular profiling of the residual disease of triple-negative breast cancers after neoadjuvant chemotherapy identifies actionable therapeutic targets. *Cancer Discov.* **4**, 232–245 (2014).

52. Almendro, V. *et al.* Inference of tumor evolution during chemotherapy by computational modeling and in situ analysis of genetic and phenotypic cellular diversity. *Cell Rep.* **6**, 514–527 (2014).

53. Savage, P. *et al.* A Targetable EGFR-Dependent Tumor-Initiating Program in Breast Cancer. *Cell Rep.* **21**, 1140–1149 (2017).

54. Chung, W. *et al.* Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat. Commun.* **8**, 1–12 (2017).

55. Oxnard, G. R. The cellular origins of drug resistance in cancer. *Nat. Med.* **22**, 232–234 (2016).

# Chapter 2.

# Kinome rewiring reveals AURKA limits PI3K-pathway inhibitor efficacy in breast cancer

## 2.1 Abstract

Dysregulation of the PI3K-AKT-mTOR signaling network is a prominent feature of breast cancers. However, clinical responses to drugs targeting this pathway have been modest, possibly because of dynamic changes in cellular signaling that drive resistance and limit drug efficacy. Using a quantitative chemoproteomics approach, we mapped kinome dynamics in response to inhibitors of this pathway and identified signaling changes that correlate with drug sensitivity. Maintenance of AURKA after drug treatment was associated with resistance in breast cancer models. Incomplete inhibition of AURKA was a common source of therapy failure, and combinations of PI3K, AKT or mTOR inhibitors with the AURKA inhibitor MLN8237 were highly synergistic and durably suppressed mTOR signaling, resulting in apoptosis and tumor regression *in vivo*. This signaling map identifies survival factors whose presence limits the efficacy of targeted therapies and reveals new drug combinations that may unlock the full potential of PI3K–AKT–mTOR pathway inhibitors in breast cancer.

## 2.2 Introduction

Mutations and aberrant signaling of the PI3K–AKT–mTOR pathway (PI3K pathway) is a prominent feature of breast cancer and many other cancer types. Genomic alterations of PI3K-pathway components including *PTEN, PIK3CA* and *AKT1* occur in over 60% of breast malignancies[1]. Despite this high prevalence, drugs targeting this pathway have demonstrated only modest responses across numerous clinical trials[2,3]. The clinical observation that most breast cancers fail to respond suggests that additional factors modulate cellular response and drive resistance. A prominent feature of this pathway is drug-induced signaling adaptation and feedback mechanisms resulting in suboptimal drug responses[4–6]. Therefore, it is likely that understanding and targeting these dynamic changes in signaling will be important for optimizing this class of agents.

In principle, the measurement of dynamic changes elicited by therapy can be used to develop novel drug combinations. Though previous efforts have focused on acute signaling changes leading to pathway reactivation and drug resistance[4,7], a systematic comparison of global signaling changes with drug efficacy has not been performed. Such an analysis may reveal survival factors whose suppression is required for drug efficacy, and hence could reveal new combinatorial strategies to enhance therapeutic responses. Previous identification of such factors has led to the understanding that drug-induced activation of apoptotic machinery[8,9] and impairment of protein synthesis[10] is required for sensitivity to a wide variety of drugs. In the context of breast cancer, multiple efforts in the field have identified mTORC1 as a survival factor

whose suppression is necessary for PI3K-pathway inhibitor sensitivity[11,12]. This observation has led to clinical trials combining PI3K and mTOR inhibitors, yet reported clinical results have yielded suboptimal outcomes as a result of increased systemic toxicity and cytostatic tumor effects[3]. Hence, there remains a pressing need to uncover new combination targets in order to improve therapeutic efficiency of PI3K-pathway inhibitors. Identifying additional survival factors will require a comprehensive understanding of signaling dynamics in response to treatment and insight as to how these dynamics contribute to drug resistance.

Little is known about global kinome rewiring in response to drug treatment, due in part to limitations in available technologies. Recently, a kinase-enrichment strategy has been developed using a chemoproteomics technique that combines kinase affinity capture with quantitative mass spectrometry (MS). This approach uses a multiplexed set of type I kinase inhibitors immobilized onto beads (multiplexed inhibitor beads, MIBs), which are used to affinity purify a diverse set of active kinases through their increased avidity for ATP compared to inactive kinases. Enriched kinases are then identified and quantified by LC–MS/MS (MIBs/MS), enabling simultaneous measurement of many endogenous kinases based on their activity state and abundance[7]. Because many drugs impinge on common pathways, and cell lines often display unique behaviors, it is possible that a quantitative map of kinase dynamics spanning multiple cell lines and drug treatments may be used to identify more general responses to drug treatment that are linked to drug sensitivity.

Here we applied the MIBs/MS approach to identify signaling changes associated with drug efficacy by mapping the kinome following exposure to targeted therapies across a panel of breast

cancer cell lines of various subtypes and genotypes. Comparison of kinome activity profiles between drug-sensitive and resistant cells allowed us to generate a kinome-response signature associated with drug sensitivity. By performing a systematic analysis of signaling dynamics following drug treatment, we identified that failure to inhibit AURKA was associated with resistance to a diverse set of targeted therapies. Further analysis revealed that inhibition of AURKA was sufficient to engender strong synergistic responses when combined with inhibitors of PI3K, AKT or mTOR. This provides an effective new framework for the unbiased identification of survival factors acting as molecular barriers to the efficacy of drugs, and we demonstrate the utility of this approach by developing rational combination strategies to enhance responses to PI3K-pathway inhibitors in breast cancer.

## 2.3 Results

### 2.3.1 Generation and analysis of a dynamic kinome signaling map

We applied an unbiased proteomic strategy to measure kinome rewiring in response to drug treatment. Kinome profiling was performed via a chemoproteomics approach using MIBs coupled with mass spectrometry (MIBs/MS). Our library of MIBs consists of a mixture of sepharose beads covalently linked to 12 kinase inhibitors ranging from moderately selective (for example, lapatinib or sorafenib) to pan-kinase inhibitors (for example, purvalanol B or staurosporine) for broad kinome coverage (Fig. 2.1a; Supplementary Fig. 2.1). Because type I kinase inhibitors preferentially bind kinases in their active conformation, kinase capture by MIBs

under the stringent binding conditions used here is a function of kinase expression, the affinity of kinases for the immobilized inhibitors, and the activation state of the kinase[13]. DMSO or drug-treated cell lysates were incubated with MIBs, and enriched kinases were eluted and quantified by LC–MS/MS using label-free quantitation (see Methods)[14]. We estimate that our current approach is able to capture roughly 35% of highly expressed kinases in a given sample (Supplementary Fig. 2.2).

We applied this strategy to a panel of breast cancer cell lines of various subtype and genotype classifications and measured kinome dynamics following treatment with a panel of targeted therapies. Cell lines were chosen to maximize transcriptional diversity and span the major subtypes of breast cancer (Supplementary Fig. 2.3). All lines harbored mutations in PI3K-pathway genes including *PIK3CA* mutant MCF7 ($ER^+/PR^+$), BT20 (receptor negative) and T47D ($ER^+/PR^+$); *PTEN*-null BT549 (receptor negative); and *HER2*-amplified SKBR3 ($HER2^+$) (Supplementary Fig. 2.4a). Cell lines were treated for 24 h with DMSO or kinase inhibitors relevant to breast cancer signaling, including the EGFR/HER2 inhibitor lapatinib (200 nM), the pan-class I PI3K inhibitor GDC-0941 (250 nM), the AKT inhibitor MK2206 (250 nM), and the MEK inhibitor PD0325901 (100 nM), and then profiled using MIBs/MS (Fig. 2.1a; Supplementary Fig. 2.4b). Together, we quantified changes across 151 kinases in total, with changes in 75 kinases present in over 75% (15/20) of the samples (Fig. 2.1b; Supplementary Dataset 2.1). Significant drug-induced changes (defined based on the $\log_2$ fold change of drug versus DMSO treatment, logFC) were detected for 99 kinases at $P < 0.001$, corresponding to 66% of kinases measured, indicating that the drugs had widespread and significant impacts on global kinome dynamics.

To assess the quality and reproducibility of the MIBs/MS data, we initially compared biological replicates of SKBR3 (HER2$^+$) cells treated with the dual EGFR/HER2 small-molecule inhibitor lapatinib. We observed a high correlation of 0.78 between replicates for identified kinases ($P = 5 \times 10^{-26}$) (Fig. 2.1c). The MIBs/MS screening strategy also accurately captured activity inhibition of direct drug targets by lapatinib, indicated by the significant decrease in levels for EGFR (logFC = −5.8, $P = 6 \times 10^{-5}$) and HER2 (−0.7, $P = 1 \times 10^{-4}$) (Fig. 2.1c). We observed a decrease in MEK1 activity upon treatment with the MEK inhibitor PD0325901 in BT549 and MCF7 cells (logFC = −1.8 and −1.2, respectively; Fig. 2.1d). We also observed indirect pathway-specific events, such as a decrease in the activity of the mTOR effector kinase RPS6KB1 when treated with either the PI3K inhibitor GDC-0941 or the AKT inhibitor MK2206 in MCF7 cells (logFC = −3.5 and −2.3, respectively; Fig. 2.1e). Comparison of observed kinome changes to previous MIBs/MS data revealed a high degree of concordance (Supplementary Fig. 2.5)[15]. These results highlight the reproducibility of the MIBs/MS approach, as well as its ability to identify direct and indirect drug targets based on reductions in both activity and abundance.

We hypothesized that the identification of shared responses across lines and drugs may lead to a more robust understanding of signaling dynamics, as opposed to changes specific to a particular drug or cell type. We therefore sought to identify changes that were generally associated with sensitivity or resistance to treatment in a drug-agnostic fashion. First, cell lines were classified as sensitive or resistant to each of the drugs in our panel on the basis of dose–response analysis (Supplementary Fig. 2.4b; Supplementary Fig. 2.6a-d). Next, fold changes for each kinase were compared between these sensitive and resistant classifications for all drugs pooled together to

identify candidate kinases whose inhibition was associated with drug sensitivity (Fig. 2.1f). This analysis revealed that suppression of 12 kinases was significantly associated with drug sensitivity ($P < 0.05$). Among the identified candidates were kinases involved in cell-cycle processes, including mitotic kinases AURKA ($P = 0.0001$) and CDK1 ($P = 0.04$), and kinases involved in interphase, including CDK4 ($P = 0.02$) and CDK2 ($P = 0.05$). Other kinases identified were involved in YAP signaling (STK4, $P = 0.01$) and WNT signaling (GSK3B, $P = 0.005$ and CSNK1E, $P = 0.02$). These results were not linked to general impairment of the cell cycle *per se*. We observed no correlation with sensitivity for other cyclin-dependent kinases (CDKs) measured in our screen such as CDK6, which is closely related to CDK4. In addition, the AURKA paralog AURKB was not significantly associated with sensitivity even though it is regulated during mitosis in a similar manner (Fig. 2.1f)[16]. We performed a similar analysis using a three-response categorization (i.e., sensitive, moderately sensitive and resistant) and found that these results were largely independent of the way sensitivity was classified (Supplementary Fig. 2.6e-g). We postulate that this drug-agnostic approach identifies changes that are general to drug sensitivity and reveals factors that may be missed by studies limited to a single-drug analyses. For example, the top candidate from our analysis, AURKA, was implicated but not found to be significantly associated with resistance or even among the top several candidates with any single drug. However, by pooling responses across all drugs it emerged as the one most associated with resistance in terms of both magnitude and significance (Supplementary Fig. 2.7). Therefore, by performing a systematic screen of signaling dynamics following drug exposure, we identified a set of specific kinases whose maintenance was associated with resistance to targeted therapies in breast cancer.

### 2.3.2 AURKA associates with PI3K and AKT inhibitor resistance

We focused our validation of molecular correlates of drug sensitivity on the PI3K pathway because of its central importance to breast cancer. We observed a significant association between maintenance of AURKA after treatment and drug resistance (Fig. 2.2a). To confirm this result, we measured molecular responses to treatment with the pan-PI3K inhibitor GDC-0941 in two sensitive (T47D and MCF7, $IC_{50}$ < 200 nM) and two new cell lines that were robustly resistant (HCC38 and MDAMB453, $IC_{50}$ > 40 μM). A critical output of the PI3K pathway is the activation of the mTORC1 complex, whose inhibition is necessary for sensitivity to PI3K inhibitors[11]. After treatment we observed suppression of mTORC1 activity only in sensitive cells, as evidenced by decreased phosphorylation of its effector protein S6 (Fig. 2.2b). Confirming our MIBs/MS data, in response to treatment we observed decreases in the abundance and autophosphorylation of AURKA in sensitive cells, whereas resistant cells maintained these levels throughout (Fig. 2.2b; Supplementary Fig. 2.9a,b). Similar results were observed using the AKT inhibitor MK2206, representing the next step in the PI3K pathway (Supplementary Fig. 2.9c-e). These results confirm that failure to suppress AURKA activity is associated with resistance to PI3K and AKT inhibition in breast cancer cells.

We next asked how AURKA is regulated in response to PI3K-pathway inhibition in drug-sensitive cells. AURKA regulates centrosome alignment, mitotic-spindle formation and chromosome segregation during mitosis, and its activity and abundance is tightly regulated[16]. We observed a robust and significant change in AURKA protein levels after 24 h in drug-sensitive cells, leading us to hypothesize that changes in transcription of *AURKA* might account for its

loss after treatment. AURKA mRNA levels were decreased in response to GDC-0941 and MK2206 when comparing drug-sensitive and resistant cell lines ($P = 2.8 \times 10^{-5}$ and $P = 0.004$, respectively; Fig. 2.2c,d). In addition, transcriptomes of MCF7 and T47D cells treated with the PI3Kα-specific inhibitor BYL719 for 24 h[17] reflected a significant reduction of *AURKA* after drug treatment in both of these BYL719-sensitive cell lines ($IC_{50} \leq 250$ nM; Supplementary Fig. 2.10a)[11,18]. Interestingly, Gene Set Enrichment Analysis (GSEA)[19] of these transcriptomes revealed that a prominent component of the response to PI3K inhibition was the suppression of genes involved in the G2/M checkpoint, including *AURKA*, suggesting that transcriptional control of this aspect of the cell cycle is a major output of the PI3K pathway (Fig. 2.2e; Supplementary Fig. 2.10b; Supplementary Dataset 2.2).

### 2.3.3 AURKA mediates survival during PI3K-pathway inhibition

We next asked if the downregulation of AURKA was functionally relevant and whether the presence of AURKA limits efficacy of PI3K-pathway directed therapies. We tested whether AURKA inhibition was sufficient to confer sensitivity to PI3K-pathway inhibitors using a combination-profiling approach to measure drug synergy across an extended panel of 13 breast cancer cell lines. We applied a dose matrix of increasing concentrations of the AURKA-specific inhibitor MLN8237 alone and in combination with a PI3K (GDC-0941), AKT (MK2206) or mTOR (RAD001) inhibitor and measured the effects on cell proliferation. To evaluate drug synergy, we (1) visualized Loewe excess values, (2) scored combination index values measuring shifts in drug potency, (3) calculated synergy scores based on Loewe excess values and (4) visualized and scored combinations using a Bliss independence model[20] (see Methods). Our

results in MCF7 cells indicated that MLN8237 in combination with GDC-0941, MK2206 or RAD001 was synergistic using all four approaches (Fig. 2.3a; Supplementary Figs. 2.11 – 2.13; Supplementary Dataset 2.3). By testing the combination of MLN8237 with GDC-0941 across the extended panel of cell lines, we found significant synergy based on the Loewe excess model in 38% of models (5/13) on the basis of a synergy score > 1, which we determined through simulation to represent a less than 5% chance of nonsynergy (i.e., FDR < 5%; Fig. 2.3b; Supplementary Fig. 2.11). We extended this analysis to drug combinations of MLN8237 with either MK2206 or RAD001 and found significant synergy in 54% and 85% of models, respectively (Fig. 2.3b; Supplementary Figs. 2.12 and 2.13). Overall, we found no significant trend toward synergy based on *PIK3CA* or *PTEN* mutational status, but did observe slightly increased synergy in receptor-positive cell lines ($ER^+$ or $HER2^+$; $P = 0.04$ for GDC-0941 and $P = 0.035$ for MK2206, based on a two-tailed *t*-test; Supplementary Dataset 2.3).

Because PI3K-pathway inhibitors are primarily cytostatic[5,21] and AURKA is known to regulate apoptosis[22], we next asked whether AURKA inhibition could enhance responses to PI3K-pathway inhibitors by inducing cytotoxic responses. Across 12 cell lines, we found that the addition of MLN8237 caused an increase in apoptotic cell death (Fig. 2.3c) that was independent of the particular dose used (Supplementary Fig. 2.14a,b). This enhancement in cell death generally occurred in conditions in which synergy was also observed (Supplementary Fig. 2.14b). We compared this response with the combination of CDK4/6 and PI3K inhibitors which are known to be synergistic[12]. Although we observed synergy between PI3K, AKT and mTOR inhibitors and the CDK4/6 inhibitor LEE011, the response was primarily cytostatic, indicating that CDK4 and CDK6 are only necessary for proliferation rather than for tumor cell

survival in the presence of PI3K-pathway inhibitors (Supplementary Fig. 2.14c-e; Supplementary Dataset 2.3). Therefore, AURKA mediates cellular survival in the context of PI3K-pathway inhibition, and because the drug combinations are synergistic in inducing apoptosis in breast cancer cells, we propose that it may be a promising companion target to enhance the efficacy of PI3K-pathway inhibitors.

### 2.3.4 MLN8237 and Everolimus (RAD001) induce cell death *in vivo*

We next evaluated the efficacy of this combination in vivo and focused on the combination of MLN8237 with the only FDA-approved inhibitor targeting this pathway in breast cancer, the mTOR inhibitor RAD001 (everolimus). Clinically, RAD001 overwhelmingly results in disease stabilization rather than regression[23]. This is reflected in vitro, wherein all lines have a high RAD001 $E_{max}$, indicating cytostatic effects. In particular, MCF7 cells have a high $E_{max}$ of 0.54 and do not display evidence of poly(ADP-ribose) polymerase (PARP) cleavage at high doses (Supplementary Fig. 2.15). To investigate whether AURKA suppression enhances the response to RAD001 treatment, we tested the combination in MCF7 orthotopic transplants. Though RAD001 or MLN8237 monotherapy only partially impaired tumor growth, the combination showed significantly greater tumor growth inhibition than either single agent alone (Fig. 2.4a). Furthermore, all animals receiving the combination therapy (9/9) showed marked tumor regression, whereas no regressions were observed with monotherapy (0/13 in total; $P = 2 \times 10^{-6}$ by Fisher's exact test; Fig. 2.4b). Post-treatment tumor specimens displayed an induction of apoptosis specific to the combination, as demonstrated by an increase in the number of TUNEL-positive cells (Fig. 2.4c,d). During the course of study, we did not observe any

significant weight loss in animals receiving the combination as compared to the RAD001 single-agent group (Supplementary Fig. 2.16), suggesting tolerability and a lack of added toxicity from co-inhibiting Aurora kinase A. Therefore, addition of MLN8237 to RAD001 treatment results in tumor regression and a strong cytotoxic response in vivo.

### 2.3.5 Co-inhibition durably suppresses mTORC1 signaling via AKT

We next turned to identification of the mechanisms driving the increased efficacy of the drug combination. Because most PI3K-pathway inhibitors (including rapamycin and RAD001) elicit feedback signals resulting in incomplete suppression of mTOR and drug resistance[11,24], we first asked whether the combination of MLN8237 enhanced the activity of RAD001 on mTOR signaling to effectors RPS6 (S6) and 4E-BP1 in vivo. Though we observed an incomplete and partial suppression of S6 in RAD001-treated MCF7 xenografts, the addition of MLN8237 resulted in a durable and complete loss of S6 in all nine tumors (Fig. 2.5a). Though RAD001 is a relatively potent inhibitor of S6, it is a weak inhibitor of 4E-BP1, and therefore only partially impairs cap-dependent protein synthesis[24]. We thus investigated the activity of phospho-4E-BP1, which can be stimulated by rapamycin treatment[24]. Although phospho-4E-BP1 levels were enhanced with RAD001 single-agent treatment, co-treatment with MLN8237 suppressed these levels back to nearly baseline (Fig. 2.5a). This surprising finding led us to ask how Aurora kinase inhibition might alter this key signaling output of mTOR. We investigated AKT activity via phosphorylation of serine 473, which activates mTOR and is catalyzed by a variety of kinases[25]. Single-agent MLN8237 reduced phospho-AKT levels in both monotherapy and combination treatments, indicating that Aurora kinases sustain mTOR levels by promoting AKT activity

(Fig. 2.5a). We next examined whether Aurora-kinase-driven maintenance of mTOR was a general feature of PI3K-pathway inhibitors. Using MCF7 cells in vitro, we observed that MLN8237 treatment impaired phospho-AKT and that the combination of MLN8237 with either GDC-0941 (targeting PI3K) or MK2206 (targeting AKT) led to robust ablation of phospho-S6 and phospho-4E-BP1 levels (Fig. 2.5b). Therefore, Aurora kinases contribute to resistance to PI3K-pathway inhibitors through the maintenance of AKT and residual mTORC1 activity. Hence, targeting this survival mechanism results in a more durable and complete repression of the PI3K pathway.

### 2.3.6 Co-inhibition unbalances pro- and antiapoptotic factors

As we observed cell death in response to these drug combinations (Figs. 2.4d and 2.5b), we next sought to elucidate how Aurora kinase mediates cell survival in response to PI3K-pathway suppression. Aurora kinases and mTOR both regulate a number of components of the intrinsic apoptosis pathway[22,26], and we hypothesized that deregulation of the balance of pro- and antiapoptotic factors may cause cell death in response to drug combinations containing MLN8237. BAX promotes apoptosis, whereas BCL2 prevents apoptosis by inhibiting the activity of BAX, and together the balance of these two proteins forms a molecular rheostat for apoptosis[27]. In MCF7 xenografts, combination treatment resulted in an increase in BAX levels and a reduction in BCL2 levels, leading to an increase in the ratio of BAX/BCL2 compared to either MLN8237 or RAD001 treatment alone (Fig. 2.5c). Furthermore, the BAX/BCL2 ratio was also increased by the addition of MLN8237 to GDC-0941, MK2206 or RAD001 in MCF7 cells in vitro, in which it was associated with the presence of cleaved PARP (Fig. 2.5b,d). Taken

together, we propose a model whereby Aurora kinase inhibitors potentiate the activity of PI3K-pathway inhibitors by enabling a durable and complete suppression of AKT/mTOR signaling and drive cell death by altering the balance of pro- and antiapoptotic factors (Fig. 2.5e).

### 2.3.7 MYC regulates AURKA downstream of the PI3K pathway

We next sought to identify factors that regulate AURKA in response to PI3K-pathway inhibition. We noted that a MYC target gene signature was among the most suppressed gene sets after treatment with BYL719, suggesting that MYC may play a significant role in regulating the transcriptional response to PI3K inhibition and therefore potentially AURKA (Fig. 2.2e). To directly define whether MYC activity is suppressed by PI3K-pathway inhibition, we transcriptionally profiled an isogenic pair of MCF10A breast epithelial cells overexpressing MYC to derive a gene signature of the top 150 most upregulated genes by MYC (Supplementary Dataset 2.4). Comparison of this signature to transcriptional changes induced by BYL719 treatment in MCF7 and T47D cells revealed that most MYC signature genes were strongly repressed during PI3K inhibition (Fig. 2.6a,b). Therefore, MYC is regulated by the PI3K pathway in these cells, likely via mTORC1-mediated translation and AKT-mediated stabilization of MYC[28–30]. AURKA was among the signature genes, and we found that MYC-overexpressing cells had an eight-fold increase in AURKA transcript levels as well as higher levels of total and phosphorylated AURKA protein (Fig. 2.6c,d). These data provide direct evidence that MYC regulates AURKA abundance and activity and suggest that both are controlled by the PI3K pathway in breast cancer.

Considering that AURKA activates AKT (Fig. 2.5b)[31,32], our results suggest a model whereby the PI3K pathway regulates the abundance of its upstream activator AURKA through the control of MYC. Hence, MYC-driven AURKA signaling may constitute a positive feedback loop that helps to continuously activate the PI3K pathway, even in the context of single-agent drug treatment. In support of this theory, we observed that MCF10A-MYC cells were more resistant to GDC-0941 and MK2206 compared to parental cells, consistent with previous reports of MYC driving resistance to inhibitors of this pathway (Fig. 2.6e,f; Supplementary Fig. 2.19a,b)[33–36]. Although MYC-expressing cells were drug resistant, they could be resensitized to GDC-0941 or MK2206 by the addition of MLN8237 until they were back to approximately the same relative $IC_{50}$ as parental cells with this combination (Fig. 2.6e,f; Supplementary Fig. 2.19a,b), indicating that AURKA is principally responsible for causing the resistance to PI3K inhibition seen as a result of MYC activation in this model.

To test this model, we asked whether MYC-driven resistance to PI3K inhibitors occurs through the maintenance of PI3K-pathway activity and if this is dependent on AURKA. GDC-0941 treatment in MCF10A cells led to a reduction in MYC and AURKA signaling, as well as phospho-S6 and phospho-4E-BP1, indicating that MYC and AURKA are regulated by the PI3K pathway (Fig. 2.6g). However, constitutive expression of MYC resulted in the maintenance of all of these factors after PI3K inhibition, suggesting that MYC also acts upstream of the PI3K pathway and can maintain its activity. Furthermore, maintenance of mTORC1 signaling by MYC overexpression was reversed by co-inhibition of AURKA, thus designating AURKA as the critical link between MYC and activation of the PI3K pathway in these cells (Fig. 2.6g). Similar results were observed using the AKT inhibitor MK2206

(Supplementary Fig. 2.19c). Taken together, our data define a novel circuit whereby the PI3K pathway regulates the abundance of its own activator through MYC-mediated transcription of AURKA (Fig. 2.6h).

## 2.4 Discussion

Through an unbiased proteomics approach to assay kinase activity, we measured dynamic changes elicited by therapy as a means to develop novel drug combinations. The systematic measurement of kinome dynamics across a diverse set of cell lines allowed us to map molecular changes associated with resistance to a variety of inhibitors, which is unique from previous approaches limited to a single drug or cell line[7,15,37]. We found a number of cases in which failure to inhibit a particular kinase was associated with drug resistance. As our proteomic screen included multiple drugs that impinge on distinct oncogenic pathways, we found it surprising that a set of common survival factors were identified. This may be due to the convergence of both the PI3K and MAPK pathways on protein synthesis[38,39]. Beyond AURKA, we identified that CDK4 suppression was associated with drug sensitivity and that the combination of CDK4 and PI3K-pathway inhibitors was synergistic, consistent with previous work[12]. Future work may determine if other candidates we identified also act as survival factors and how they might do so.

We show that the expression of AURKA limits the efficacy of PI3K-pathway-targeted therapy and thus represents a new vulnerability that can be used to enhance therapeutic responses to this

class of drugs. By investigating AURKA regulation, we found that the reduction in AURKA abundance in drug-sensitive cells appears to be the result of transcriptional control by MYC, which is in turn regulated by the PI3K pathway. MYC has been shown to regulate AURKA transcription in multiple tumor types[40–42] and has independently been associated with resistance to PI3K inhibitors, which may be clinically relevant but remains mechanistically ambiguous[33–36]. Here we show that MYC-driven AURKA activation results in maintenance of the PI3K pathway despite PI3K inhibitor treatment, resulting in drug resistance. Future work may gauge the relative importance of AURKA relative to other outputs of MYC in driving resistance to PI3K inhibitors.

Maintenance of AURKA was sufficient to confer drug resistance in a variety of cell lines, as evident by the widespread drug synergy observed. We show that in response to treatment with PI3K-pathway inhibitors, Aurora kinase maintains the activation of AKT and drives residual mTOR activity. Co-inhibition of the PI3K pathway and AURKA with MLN8237 fully blocks this residual mTOR activity, resulting in cell death. These findings also highlight the importance of AKT activation through serine 473 as a route of drug resistance. Because a number of kinases have been shown to operate at this site, including mTORC2[25], it remains unclear whether Aurora kinases act on this site directly or indirectly. These studies elaborate a positive feedback loop whereby the PI3K pathway promotes the expression of AURKA, which in turn activates the pathway via AKT. One feature of such a positive-feedback loop is the creation of switch-like outputs resulting in heightened stability and resistance to perturbation[43]. We postulate that such loops are common and may lead to the resiliency and adaptation that is a hallmark of the PI3K pathway and a major cause of the challenges in targeting it therapeutically. Delineating such

loops may be an important strategy in identifying effective drug combinations. As a case in point, we show that eliminating this positive-feedback loop by blocking AURKA renders cells more sensitive to PI3K inhibitors.

Our findings reveal that the combination of Aurora kinase inhibitors and PI3K-pathway inhibitors is synergistic and could be a promising clinical strategy to enhance the treatment response in breast cancer. These data are consistent with observations made in other settings[44–46]. Clinical data of PI3K and mTOR inhibitors have shown only modest benefit in breast cancers, at best resulting in short-term disease stabilization in patients[23,47]. Consistent with these clinical observations, most inhibitors in this class cause only a proliferative arrest in vitro[5,21] and it has been proposed that combinations that induce apoptosis may be used to enhance responses[48]. In contrast to cytostatic combinations with the CDK4/6 inhibitor (i.e., synthetic sickness), we found that combinations with Aurora kinase inhibitors were synergistic and potently induced cell death. As clinical trials testing CDK4/6 inhibitor combinations are ongoing, it remains to be seen what impact this distinction will play on patient responses. These results warrant an expanded analysis of combinations with AURKA inhibitors in additional patient-derived models of breast cancer and other cancer types. Tested as monotherapy, Aurora kinase inhibitors have reached phase 3 clinical trials for lymphoma with manageable toxicities but limited efficacy[49]. Given that the most common adverse events of PI3K-pathway inhibition are hyperglycemia, rash and gastrointestinal toxicity, and that those of Aurora kinase inhibition are primarily neutropenia, we are encouraged that the nonoverlapping toxicity profile between the two agents may be tolerated in patients as they were in our in vivo studies. As single-agent responses to both

PI3K-pathway and Aurora kinase inhibitors have been modest, these findings may unlock the full potential of these agents in realizing a clinical benefit.

## 2.5 Methods

### 2.5.1 Breast cancer cell lines and reagents

BT549 and SKBR3 cells were obtained from the UCSF Cell Culture Facility. BT20, BT474, HCC1428, HCC38, LY2, MCF7, MDAMB231, MDAMB453, T47D, SUM52PE, and ZR75B cell lines were obtained from the American Type Culture Collection (ATCC). Cell lines used for proteomic profiling and molecular analyses were authenticated by STR analysis. Lines were grown according to published protocols[50] except for SKBR3, which was cultured using RPMI media supplemented with 10% FBS and 1% pen/strep. All cell lines tested negative for mycoplasma contamination. Drugs used for cell culture experiments in this study were purchased from Selleck Chemicals (GDC-0941, MK2206, PD0325901, lapatinib, MLN8237, and LEE011) and LC Laboratories (RAD001).

### 2.5.2 Multiplex inhibitor bead (MIB) analysis

Multiplexed inhibitor bead enrichment and MS analysis (MIBs/MS) were performed as described previously[14]. In summary, a selection of bait compounds were purchased or synthesized and immobilized on sepharose using standard peptide coupling chemistry. The following compounds were purchased commercially: bisindolylmaleimide X (Enzo Life Sciences);

SB202190, staurosporine (LC Labs); purvalanol B (Tocris); lapatinib, crizotinib, dasatinib (Selleckchem). When not commercially available without modification, linkable versions of previously described compounds were synthesized based on prior methods: VI-16832[51,52], Akti-46[53], PP-hydroxyl[54], sorafenib[55], and JG-4[56] with minor adjustments made for synthetic tractability. After initial pilot syntheses and validation, compounds were synthesized by Pharmaron, Inc. Louisville KY. Couplings were performed overnight at room temperature (20–25 ˚C) on a rotator. Beads and compounds were mixed in 1:1 of dimethyl formamide:ethanol with 0.1 M 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide.

After 24-h treatment with drug or DMSO, cell lysates were diluted in binding buffer with 1 mol/L NaCl, and kinase enrichment was performed using gravity chromatography following preclearing. After washing, the bound kinases were eluted with SDS followed by extraction/precipitation, tryptic digest and desalting. Liquid chromatography–tandem mass spectrometry (LC–MS/MS) was performed on a Velos Orbitrap (Thermo Scientific) with in-line high-performance liquid chromatography (HPLC) using an EASY-spray column (Thermo Scientific). Peptide identifications were made using ProteinProspector (v5.10.10) and input into Skyline for label-free quantification[57].

Peptide quantification data were pre-processed before analysis with MSstats v2.3.3[58]. First, library peptides and peptides that map to nonkinase proteins were removed. Kinase peptide peak area values were $\log_2$-transformed and quantile-normalized to correct for variation between replicates. Finally, peptides that mapped to multiple kinases were removed, as well as peptides that were entirely missing in one or more conditions. For each kinase, the $\log_2$ ratio of each

drug-treated condition to the DMSO control was estimated using the mixed-effects regression model in MSstats.

### 2.5.3 Drug combination studies

Cell lines were seeded in 384-well assay plates at a density of 1,000 cells/well in a total volume of 40 µL/well and incubated at 37 ˚C, 5% $CO_2$ overnight. Dose matrices were assembled containing six-point, four-fold serial dilutions from the top concentration for each agent on the *x*- and *y*-axes. Following 72 h of drug exposure, proliferation and cell death was measured by staining with Hoescht (Life Technologies) nuclear dye and YO-PRO-1 (Life Technologies), respectively, and analyzed using a Thermo CellInsight High Content microscope. Raw phenotype measurements from each treated well were normalized to the median of vehicle-treated control wells and examined for synergistic effects between both compounds.

To evaluate drug combinations, we used a Loewe model of drug additivity and calculated a synergy score. First, we fit a sigmoidal function to each of the single-agent responses. Next, we calculated the expected inhibition for each combination using the Loewe additivity model[20]. The synergy score *S* was calculated as previously defined[59] as a positive-gated inhibition-weighted volume over of Loewe additivity:

$$S = lnf_X lnf_Y \sum \max(0, I_{data}) \max(0, I_{data} - I_{Loewe})$$

Where $f_X$ and $f_Y$ are the dilution factors used for compounds *X* and *Y*, respectively, $I_{data}$ is the matrix of inhibition data at this dilution factor, and $I_{Loewe}$ is the expected inhibition according to Loewe additivity. Synergy score calculations were also derived using Bliss independence[20], based

on a model in which drugs act independently of each other. $CI_{50}$ values for equal-dose combinations were calculated as previously defined[20]:

$$CI_{50} = \frac{(D)_1}{(D_{50})_1} + \frac{(D)_2}{(D_{50})_2}$$

Where $(D)_1$ and $(D)_2$ are the given doses of the two drugs, and $(D_{50})_1$ and $(D_{50})_2$ are the $IC_{50}$ values for each drug as a single agent.

To determine a cutoff for the synergy score, we simulated the distribution of scores generated by an additive drug combination. We generated two hypothetical compounds by sampling random shape parameters for their dose–response functions and calculated the expected Loewe model of the combination. We then added normally distributed noise to the model with variance estimated from our experimental data and calculated the resulting synergy score. This process was repeated 100,000 times to simulate the distribution of synergy scores for different additive combinations. The 95$^{th}$ percentile of this distribution was 0.91 and so we conservatively identified combinations with $S \geq 1$ as synergistic.

### 2.5.4 Western blotting and antibodies

Proteins were extracted using RIPA buffer (50 mM Tris–HCl pH 7.5, 150 mM NaCl, 0.1% sodium deoxycholate, 0.1% SDS, 1 mM EDTA pH 8.0, 1% NP-40) containing proteinase (Roche) and phosphatase (Roche) inhibitor cocktails. Samples were resolved using 4–12% SDS–PAGE gels (Life Technologies) and transferred to PVDF membranes (Millipore). Membranes were probed overnight on a 4 °C shaker with primary antibodies (1:1,000 dilution unless indicated) recognizing the following proteins: p-AKT (Ser473) (9271, Cell Signaling), AKT

(4691, Cell Signaling), p-S6 (Ser240/244) (5364, Cell Signaling, 1:20,000), p-4E-BP1 (Thr37/46) (2855, Cell Signaling), p-AURKA (Thr288) (3079, Cell Signaling), AURKA (4718, Cell Signaling), Cleaved PARP (Asp214) (9541, Cell Signaling), BCL2 (2870, Cell Signaling), BAX (2772, Cell Signaling), MYC (ab32072, Abcam), and β-actin (3700, Cell Signaling).

### 2.5.5 Mouse xenograft studies

All animal studies were conducted in compliance with all relevant ethical regulations set forth by the UCSF Institutional Animal Care and Use Committee (IACUC). 4-week old immunocompromised NOD/SCID female mice were purchased from Taconic Biosciences, and MCF7 cells used for in vivo transplant were obtained from the UCSF Preclinical Therapeutics Core. Xenograft tumors were initiated in the cleared mammary fat pads of mice bearing slow release estrogen pellets (Innovative Research of America) by orthotopic injection of $1 \times 10^6$ MCF7 cells in a 1:1 mixture of serum-free medium and Matrigel (BD Biosciences). When tumors reached ≥ 1 cm in any direction via electronic caliper measurements, mice were randomized into cohort groups and treatment was initiated.

Treatment arms received either vehicle (1:1 mixture of single-agent diluents), RAD001 formulated as a microemulsion (2 mg/kg/q; 30% propylene glycol, 5% Tween 80), MLN8237 (10 mg/kg/q; 10% 2-hydroxypropyl-β-cyclodextrin, 1% sodium bicarbonate), or the combination daily via oral gavage. Animals were monitored daily for evidence of toxicity, including weight and skin effects, and changes in tumor size ($mm^3$) through bidirectional measurements of perpendicular diameters using electronic calipers, and calculated as $V = 1/2\,(length \times width^2)$.

Mice were sacrificed after 15 d of treatment, following which tumors were excised and a portion of the tissue fixed in 4% paraformaldehyde. The remaining tumor tissue was flash-frozen in liquid nitrogen.

### 2.5.6 Immunohistochemical analysis

PFA-fixed tumor samples were paraffin-embedded, and immunohistochemical staining of tissue sections was performed. TUNEL staining was carried out using the ApopTag Peroxidase In situ Apoptosis Detection Kit (Millipore) according to the manufacturer's instructions ($n = 15$ data points per group; five high-powered (20×) fields analyzed from separate areas of each tumor from 3 mice per experimental group). Stained slides were digitized using the Leica DMi1 Microscope (Leica Microsystems) with a 20× objective. Images were scored as the number of TUNEL-positive cells per captured field, and quantification was performed in a manner that was blinded to treatment group.

### 2.5.7 Real-time PCR

RNA was isolated according to the manufacturer's instructions (TRIzol, Life Technologies). One microgram of total RNA from each sample was subjected to first-strand cDNA synthesis according to the manufacturer's recommendations (Promega). Quantitative PCR was performed on a CFX96 Real-Time PCR detection system with a PrimeTime Gene Expression Master Mix (IDT technology) according to the manufacturer's protocol. AURKA was amplified with the following primers: 5′-AGTTGGCAAACGCTCTGTCT-3′ (forward primer) and 5′-GTGCCACACATTGTGGTTCT-3′ (reverse primer). RPL13A was used as an endogenous

control with the following primers: 5′-CGGATTTGGTCGTATTGG-3′ (forward primer) and 5′-TCCTGGAAGATGGTGATG-3′ (reverse primer). The cycling conditions for AURKA and RPL13A were as follows: one cycle at 95 °C for 3 min; 40 cycles of 95 °C for 15 s, and 60 °C for 60 s. The specificity of the PCR amplification was validated by the presence of a single peak in the melting curve analyses.

### 2.5.8 Gene Set Enrichment Analyses (GSEA)

Gene set enrichment analysis (GSEA) of hallmark cancer gene signatures in the Molecular Signatures Database (MSigDB v6.0) was performed using GSEA v3.0 software (http://www.broadinstitute.org/gsea/)[19] under the following parameters: permutation, phenotype; metric, Signal2Noise; scoring scheme, weighted; and number of permutations, 1,000. Gene sets were considered significantly enriched following a nominal $P < 0.05$ and FDR < 0.25 cutoff.

### 2.5.9 Statistical analysis

Data are expressed as means ± s.d., unless otherwise indicated. Statistical analyses were performed using GraphPad Prism 6 (v6.0 g) and R (v3.32). Two-tailed Student *t*- tests (with unequal variance) were used in all comparisons unless otherwise noted. $P < 0.05$ was considered statistically significant throughout the study.

### 2.5.10 Reporting Summary

Further information on experimental design is available in the Nature Research Reporting Summary (https://www.nature.com/articles/s41589-018-0081-9#MOESM2) linked to this article.

### 2.5.11 Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files. The raw mass spectrometry data is accessible via http://prospector2.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer under the search key: *lixlgarvea*.

## 2.6 Acknowledgements

The authors would like to thank members of the Bandyopadhyay laboratory for helpful discussions and technical assistance. We also thank A. Beardsley, E. Markegard, D. Ruggero and W. Weiss for helpful discussions and reagents. This work was supported in part by NCI U01CA168370 (S.B.), NIGMS R01GM107671 (S.B.), NCI R01CA170447 (A.G.), Prospect Creek Foundation (S.B., A.G.), OHSU Pilot Project Funding (S.B., J.K.), American Cancer Society Postdoctoral Fellowship (J.D.G.), the Gazarian Foundation (A.G.), DOD W81XWH-12-1-0272 and DOD W81XWH-16-1-0603 (A.G.).

## 2.7 Author contributions

H.J.D., J.T.W., K.M.S, J.K., J.D.G., and S.B. contributed toward study conceptualization. H.J.D., J.T.W., and J.D.G. performed data analyses supporting the study. H.J.D. designed and performed the majority of experiments. N.B. assisted with samples for initial MIBs/MS profiling. R.S.L. J.T.W., and J.D.G. provided technical advice and guided the interpretation of mass spectrometry data from MIB/MS profiling. R.C. and O.M. assisted with animal studies, and K.N.S. helped with additional experiments. H.J.D. and S.B. composed the original draft, and all authors contributed to manuscript finalization. S.B. and A.G. supervised the study.

## 2.8 Figures

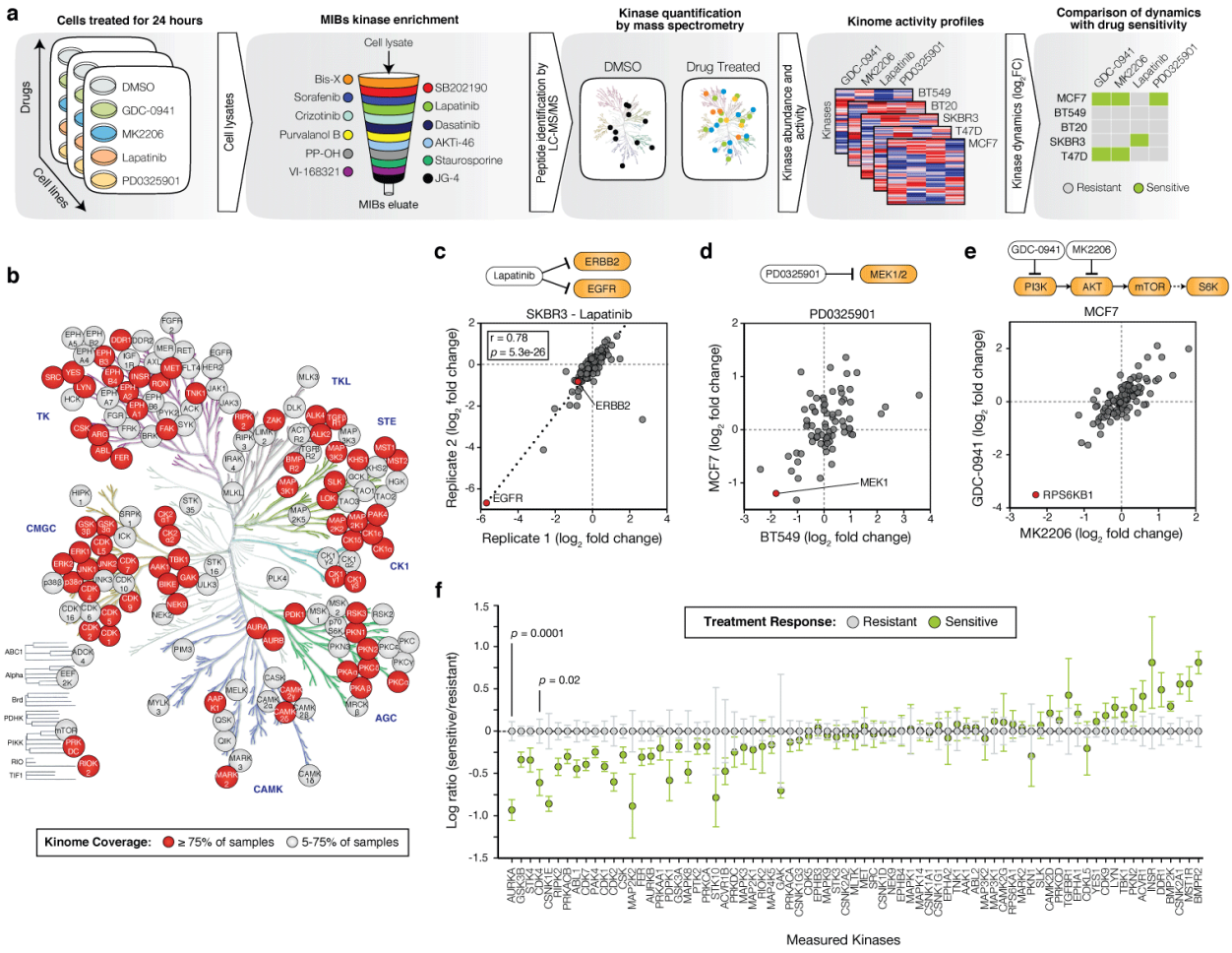



**Figure 2.1 Measurement of kinome dynamics to identify correlates of drug sensitivity.** (**a**) Schematic of the approach using multiplexed inhibitor beads followed by MS (MIBs/MS). Sample lysates are passed through a column containing the indicated kinase inhibitors covalently linked to beads. After washing, bound proteins are eluted, trypsin digested and quantified through label-free MS. (**b**), Human kinome tree annotated with kinases identified in this study, colored based on the percentage of total samples, whereby each particular kinase could be quantified. (**c**) Comparison of activity ratios between biological replicates for 122 kinases, expressed as a log ratio of measurements from SKBR3 cells treated with 200 nM lapatinib for 24 h versus DMSO. Pearson correlation and *P* value are shown. (**d**) Comparison of kinase activity ratios in BT549 and MCF7 cells treated with 100 nM PD0325901 versus DMSO. Data represent 75 kinases with one outlier kinase (GAK, BT549 $\log_2$ fold change 8.3) removed. (**e**) Comparison of activity ratios for 70 kinases measured from MCF7 cells treated with either 250 nM MK2206 or GDC-0941 compared to DMSO. (**f**) Categorical analysis of kinome dynamics occurring in drug-sensitive treatment responses ($n = 6$) versus resistant treatment responses ($n = 14$) for all drugs pooled together. For visualization purposes, each kinase was centered on the mean of resistant samples. Data are shown for 75 kinases, which could be

measured in > 75% of samples. All drug treatments are 24 h. Error bars are mean ± s.e.m., and *P* values were calculated using a two-sided *t*-test.

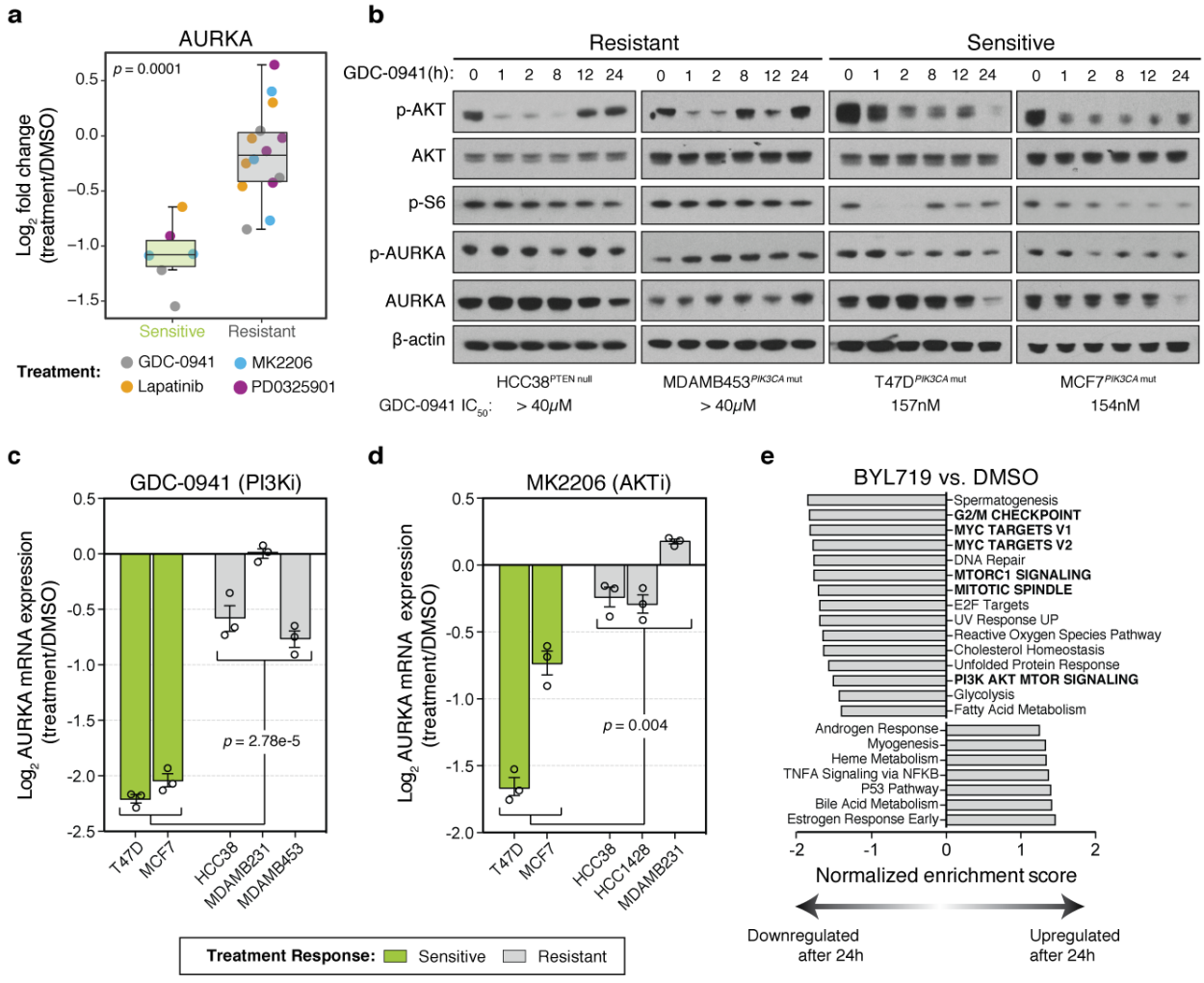



**Figure 2.2 Maintenance of AURKA is associated with resistance to PI3K inhibition.** (**a**) Changes in activity of AURKA as measured by MIBs in drug-sensitive versus drug-resistant treatment responses after 24 h of exposure to the indicated compounds. Each point reflects a single cell line and drug treatment ($n = 20$ biologically independent samples). Box-and-whisker plots show median (centerline), upper/lower quartiles (box limits), and whiskers spanning the interquartile range from 25–75 percentiles. *P* value was calculated using a two-sided *t*-test. (**b**) Western blot showing PI3K and AURKA signaling in GDC-0941-resistant and GDC-0941-sensitive cell lines. Protein lysates from cells treated with 1 μM GDC-0941 were extracted at different time points, separated by SDS–PAGE, and analyzed by immunoblot with the indicated antibodies. Image represents $n = 3$ independent experiments (full blots shown in Supplementary Fig. 2.8). (**c**,**d**) Log ratio expression values of AURKA mRNA measured by RT-PCR from the indicated cell lines treated with 1 μM of GDC-0941 (**c**) or 1 μM MK2206 (**d**) for 24 h compared to DMSO treatment. Data represents $n = 3$ biological replicates. Error bars are mean ± s.e.m., and *P* values were calculated using one-way ANOVA. (**e**) Gene Set Enrichment Analysis (GSEA) of top gene sets significantly upregulated or downregulated after 24 h in response to 1 μM BYL719 treatment in MCF7 and T47D cells compared to DMSO. Data in **e** were based on transcriptomic data from Bosch et al.[17]. PI3Ki, PI3K inhibitor; AKTi, AKT inhibitor.

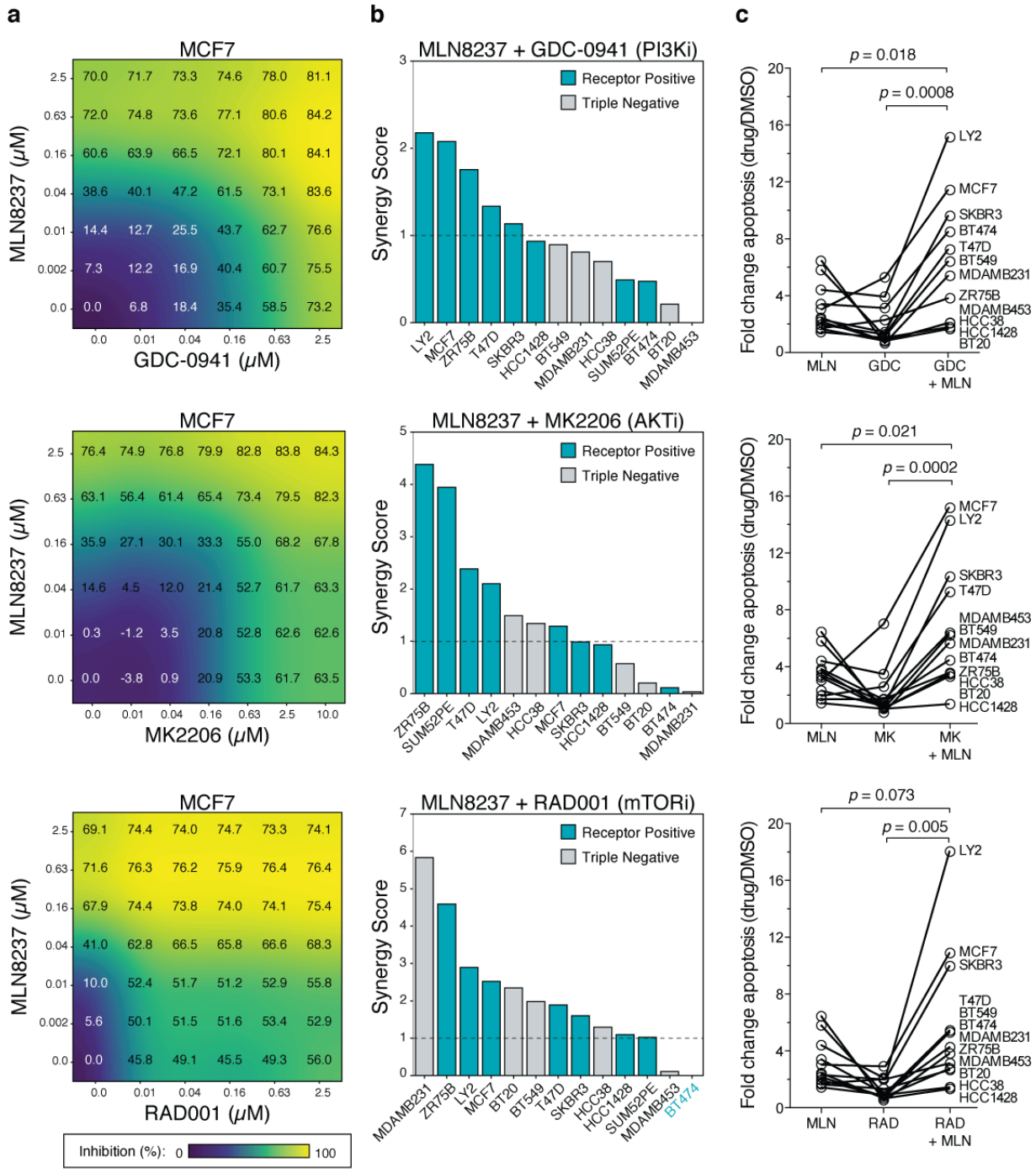



**Figure 2.3 AURKA suppression enhances sensitivity and drives cell death in response to PI3K-pathway inhibitors in breast cancer cell lines.** (**a**) A dose matrix of GDC-0941 (PI3Ki), MK2206 (AKTi), or RAD001 (mTORi) in combination with the AURKA inhibitor MLN8237 in MCF7 cells. Cell proliferation was assessed after 72 h. Percent growth inhibition at each dose is shown. (**b**) Synergy scores based on a Loewe excess inhibition model across 13 breast cancer cell lines that were treated with the indicated combinations using an escalating dose matrix for 72 h. Dashed line indicates a 5% FDR cutoff to define synergistic combinations

(see Methods). (**c**) Cell lines were treated with 625 nM of the indicated single agents or those combined together for 72 h, and apoptosis was measured by YO-PRO1 positivity. Data represents $n = 4$ biologically independent samples. Error bars are mean ± s.d., and $P$ values were calculated using a two-sided $t$-test. PI3Ki, PI3K inhibitor; AKTi, AKT inhibitor; mTORi, mTOR inhibitor.

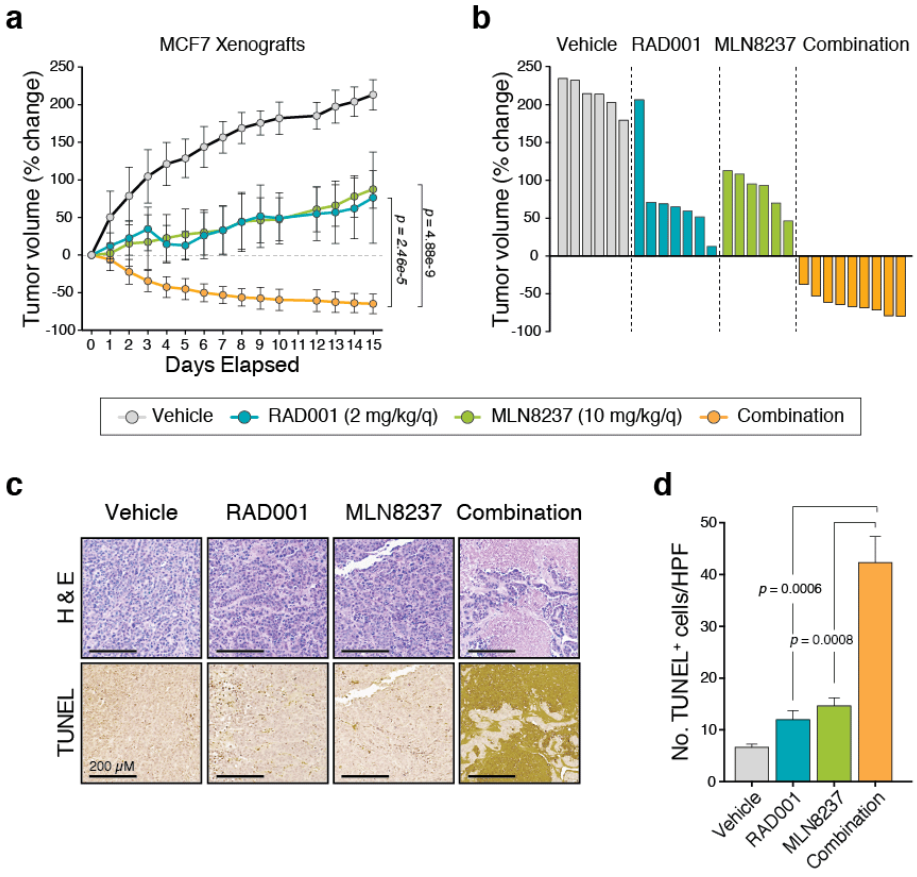



**Figure 2.4 The Aurora kinase inhibitor MLN8237 enhances sensitivity to everolimus (RAD001) and induces cell death *in vivo*.** (**a**) MCF7 orthotopic xenograft tumors were treated with vehicle ($n = 6$ biologically independent mice), RAD001 (2 mg/kg/day; $n = 7$ biologically independent mice), MLN8237 (10 mg/kg/day, $n = 6$ biologically independent mice) or with the combination of the two single agents ($n = 9$ biologically independent mice) via oral gavage, daily, over 15 d. The percentage change in tumor volume was calculated for each animal from baseline. (**b**) Individual tumor profiles compared to baseline for each tumor treated with vehicle ($n = 6$ biologically independent mice), RAD001 ($n = 7$ biologically independent mice), MLN8237 ($n = 6$ biologically independent mice) or the combination ($n = 9$ biologically independent mice) over a 15-d period. (**c**) Representative images of tumor tissue extracted from mice after 15 d of treatment with the indicated agents and stained for H&E and TUNEL. Images are shown using a 10× objective. Scale bars, 200 μm. (**d**) Quantification of the number of TUNEL$^+$ cells/field from TUNEL staining of MCF7 tumors following 15 d of treatment. Data is an average of five high-powered (20×) fields analyzed per tumor and are representative of $n = 3$ biologically independent animals. In all graphs, error bars are mean ± s.d., and *P* values were calculated using a two-sided *t*-test.

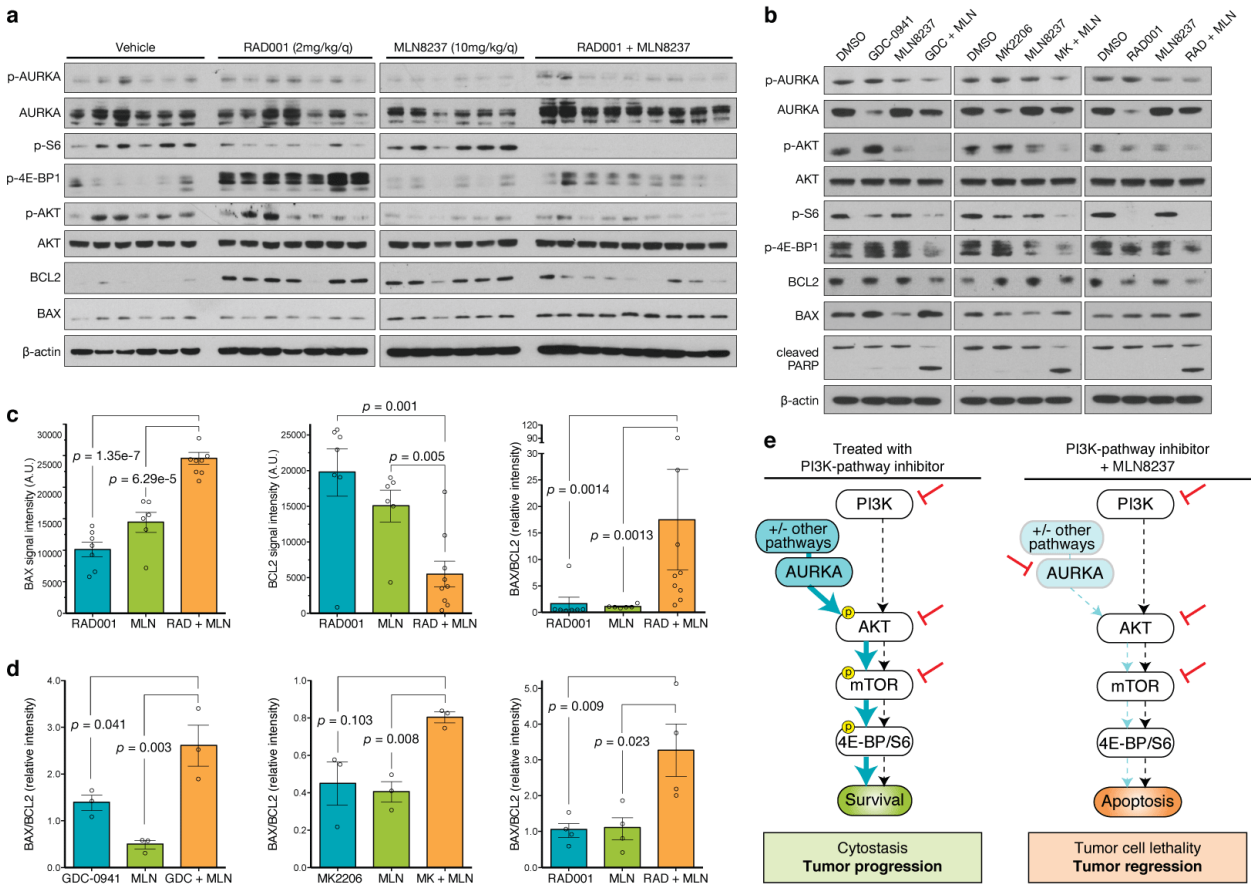



**Figure 2.5 Aurora kinase co-inhibition durably suppresses mTORC1 signaling and alters the BAX/BCL2 ratio.** (**a**) MCF7 orthotopic xenografts were treated with vehicle ($n = 6$ biologically independent mice), RAD001 (2 mg/kg/day; $n = 7$ biologically independent mice), MLN8237 (10 mg/kg/day; $n = 6$ biologically independent mice) or the combination of the two single agents ($n = 9$ biologically independent mice) for 15 d, at which point tumors were harvested and snap frozen. Western blot of protein lysates from individual tumors were probed with the indicated antibodies (full blots shown in Supplementary Fig. 2.17a). (**b**) MCF7 cells were treated with either 250 nM GDC-0941, 250 nM MK2206, 5 nM RAD001, 100 nM MLN8237 or the indicated combinations for 24 h, and protein lysates were subjected to immunoblot using the indicated antibodies. Representative image from $n = 3$ independent experiments (full blots shown in Supplementary Fig. 2.17b). (**c**) BAX, BCL2 and BAX/BCL2 ratio in MCF7 orthotopic xenografts treated for 15 d with the indicated drugs based on quantification of western blot images (RAD001, $n = 7$; MLN8237, $n = 6$; combination, $n = 9$ biologically independent mice analyzed). (**d**) BAX/BCL2 ratio in MCF7 cells treated for 24 h with the indicated drugs based on quantification of western blot images from $n = 3$ independent experiments. (**e**) Proposed model of the mechanism of Aurora kinase inhibitor synergy. De novo resistance to single-agent inhibition of PI3K, AKT or mTOR is due to incomplete suppression of the pathway because of Aurora kinase signaling, which activates AKT. Drug combinations that simultaneously inhibit the PI3K pathway and block Aurora kinase signaling completely suppress mTOR signaling to 4E-BP1 and S6, resulting in tumor cell death. In all graphs, error bars are mean ± s.e.m., and *P* values were calculated using a two-sided *t*-test.

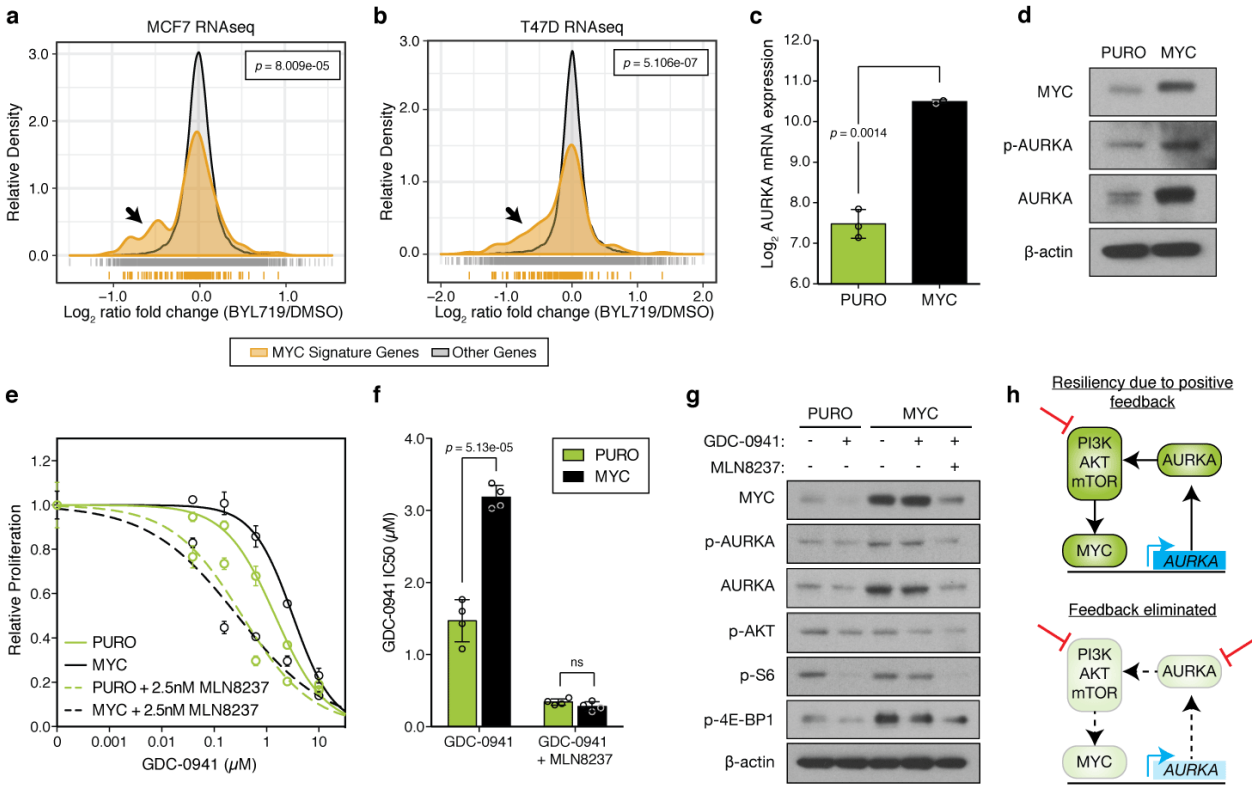
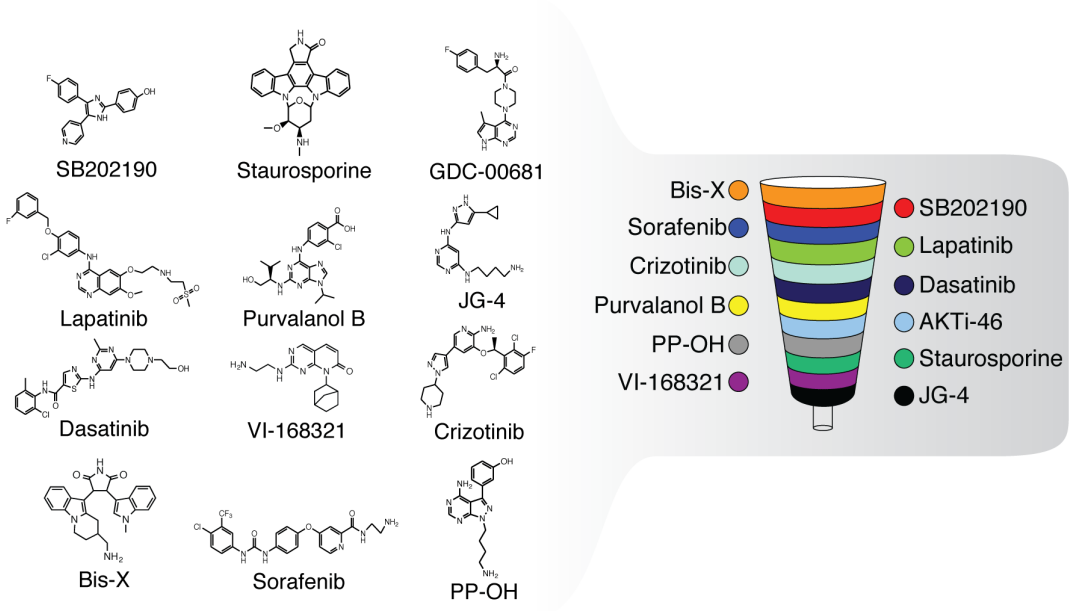
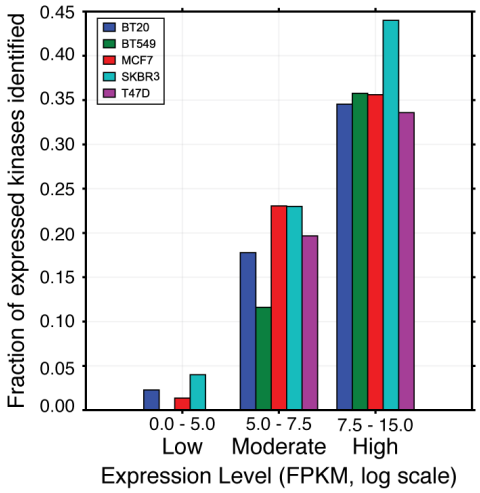



**Figure 2.6 AURKA transcription is regulated by MYC downstream of the PI3K pathway.** (**a**,**b**) Histogram of normalized gene expression of all 150 genes in the MYC gene signature compared to genes not in this signature for MCF7 (**a**) or T47D (**b**) cells treated with 1 μM BYL719 or DMSO for 24 h. BYL719 treatment data were obtained from Bosch, et al.[17] and *P* values determined by Kolmogorov–Smirnov test. Black arrows highlight differences in distribution of MYC signature genes. (**c**) Relative levels of AURKA mRNA in an isogenic pair of control (PURO, $n = 3$ independent samples) or MYC-expressing MCF10A cells ($n = 2$ independent samples) measured by RT-PCR. (**d**) Immunoblot of protein lysates from PURO or MYC cells representative of $n = 3$ independent experiments with similar results (full blots shown in Supplementary Fig. 2.18a). (**e**) Proliferation of control or MYC cells in response to GDC-0941 alone or in combination with 2.5 nM MLN8237. Combinations were normalized to MLN8237 alone. Data represent $n = 4$ biologically independent samples. (**f**) $IC_{50}$ analysis of dose–response curves shown in **e** from $n = 4$ independent samples. (**g**) Immunoblot of lysates from control and MYC MCF10A cells treated with 1 μM GDC-0941 or 100 nM MLN8237 for 24 h. Representative image of $n = 3$ independent experiments with similar results (full blots shown in Supplementary Fig. 2.18b). (**h**) Proposed model of positive feedback loop between the PI3K pathway, MYC and AURKA. In all graphs, error bars are mean ± s.d., and *P* values were calculated using a two-sided *t*-test, unless otherwise indicated.

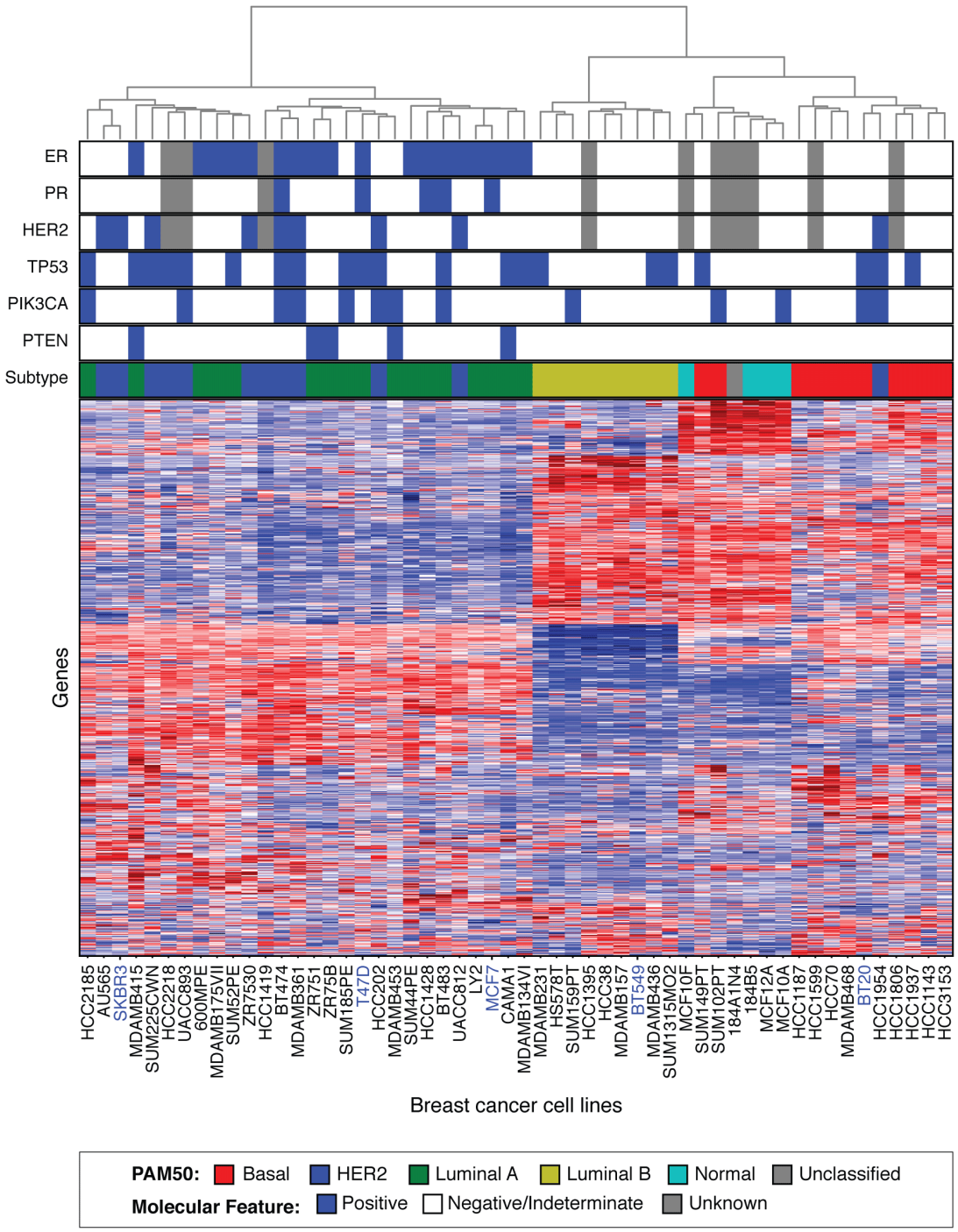
## 2.9 Supplementary Figures




**Supplementary Figure 2.1 Structure of drugs bound to multiplex inhibitor bead column.** Displayed are the inhibitors that were used for the generation of inhibitor-based beads that were assembled for the MIB kinome enrichment.




**Supplementary Figure 2.2 Dependence of MIBs data on gene expression.** Baseline RNAseq data from each listed cell line was obtained and kinases classified into three categories into high, moderate and low expression. For each cell line, the fraction of kinases expressed at each level that were captured and quantified in the MIBs dataset are shown. RNAseq data downloaded from www.synapse.org/#!Synapse:syn2346643/wiki/62255.




**Supplementary Figure 2.3 Diversity analysis of breast cancer cell lines.** Hierarchical clustering of transcriptional profiles across a panel of breast cancer cell lines. The top 1,000 most differentially expressed genes were used for clustering. Cell lines annotated based on receptor, mutation and subtype status. Cell lines used for MIBs profiling highlighted in blue. RNAseq data downloaded from www.synapse.org/#!Synapse:syn2346643/wiki/62255.

**a**

| Cell Line | Subtype | Receptor Status | *HER2* | *TP53* | *PIK3CA* | *PTEN* |
|---|---|---|---|---|---|---|
| MCF7 | Luminal | $ER^+/PR^+$ | WT | WT | MUT | WT |
| BT549 | Claudin Low | TNBC | WT | MUT | WT | Null |
| BT20 | Basal A | TNBC | WT | MUT | MUT | WT |
| SKBR3 | HER2 | $HER2^+$ | AMP | WT | WT | WT |
| T47D | Luminal | $ER^+/PR^+$ | WT | MUT | MUT | WT |

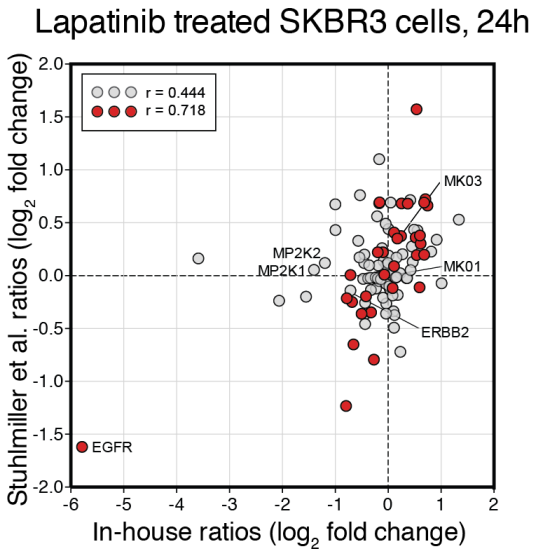
**b**

**Drug Sensitivity (μM)**

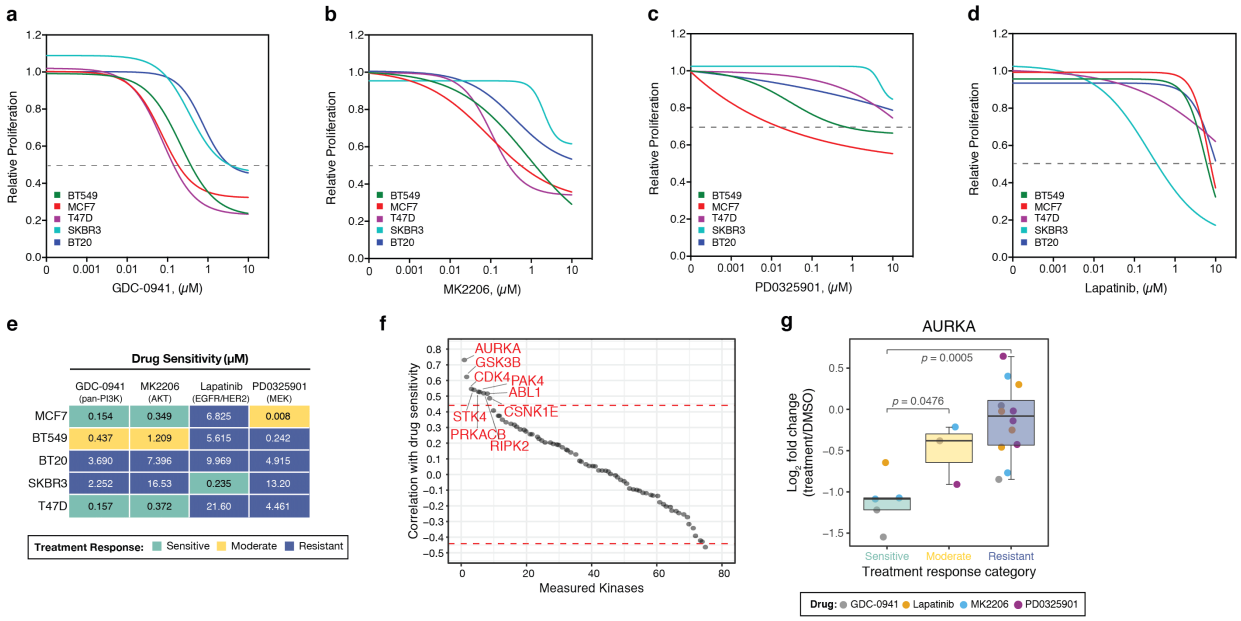
| | GDC-0941 (pan-PI3K) | MK2206 (AKT) | Lapatinib (EGFR/HER2) | PD0325901 (MEK) |
|---|---|---|---|---|
| MCF7 | 0.154 | 0.349 | 6.825 | 0.008 |
| BT549 | 0.437 | 1.209 | 5.615 | 0.242 |
| BT20 | 3.690 | 7.396 | 9.969 | 4.915 |
| SKBR3 | 2.252 | 16.53 | 0.235 | 13.20 |
| T47D | 0.157 | 0.372 | 21.60 | 4.461 |

**Treatment Response:**
Resistant
Sensitive

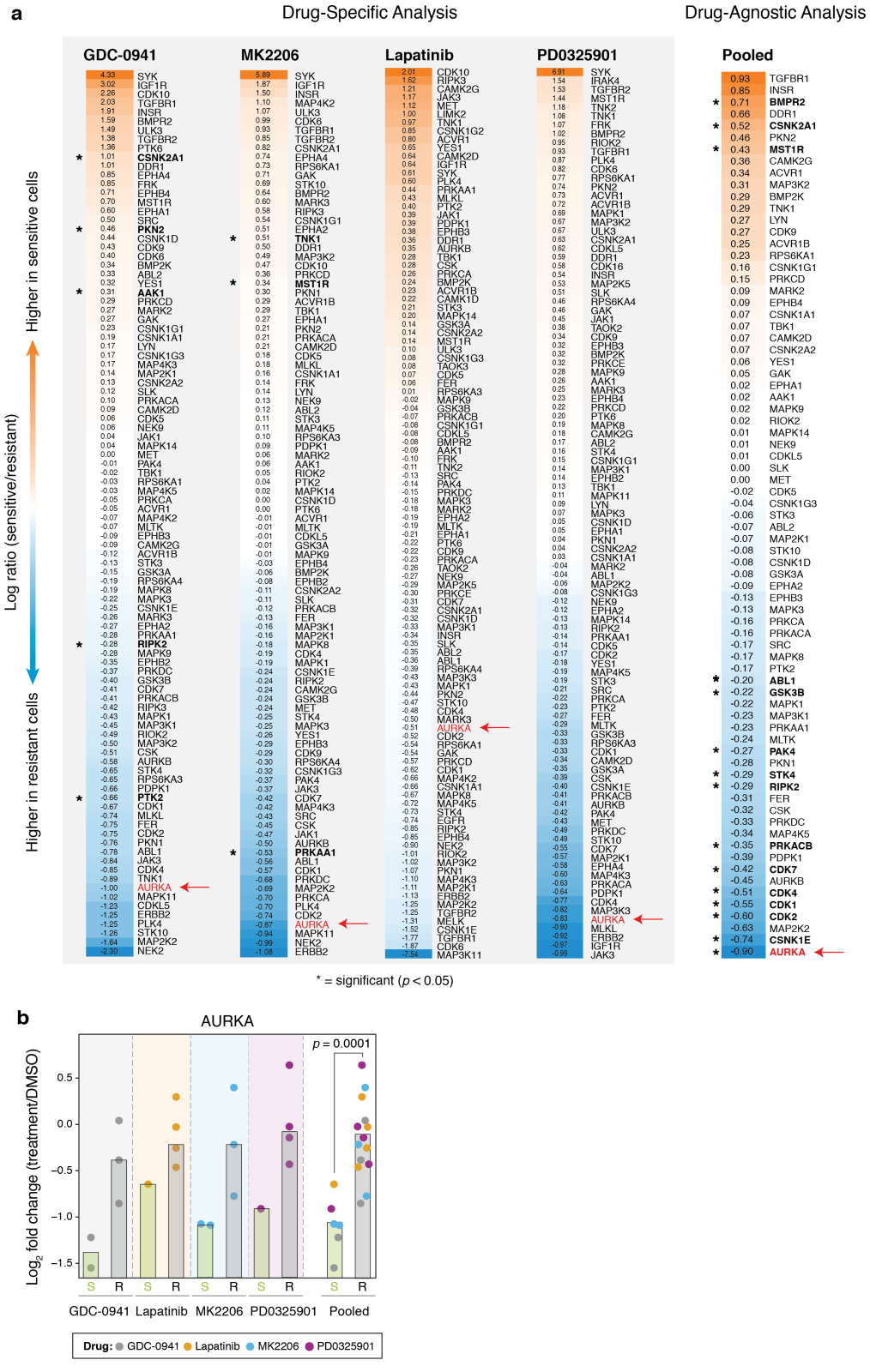
**Supplementary Figure 2.4 Characteristics of breast cancer cell lines used for kinome profiling.** (**a**) Table of breast cancer cell lines used in this study with the indicated molecular subtypes and mutational status. (**b**) Drug sensitivity of cell lines indicated by $IC_{50}$ and categorized into relative drug sensitivity or resistance. Because $IC_{50}$ was not reached for PD0325901, $IC_{25}$ is shown.




**Supplementary Figure 2.5 Drug-specific analysis of kinome changes highlight commonalities with previous reports.** Observed kinome changes (shown as kinase log ratios) for Lapatinib treated SKBR3 cells from in-house MIBs/MS screening data in comparison to kinases measured from published dataset (Stuhlmiller, et al.)[15]. Kinases reported as having significant changes in activity after treatment from the Stuhlmiller, et al. study are depicted in red. Correlation computed over $n$ = 31 kinases for red and $n$ = 104 kinases for gray.

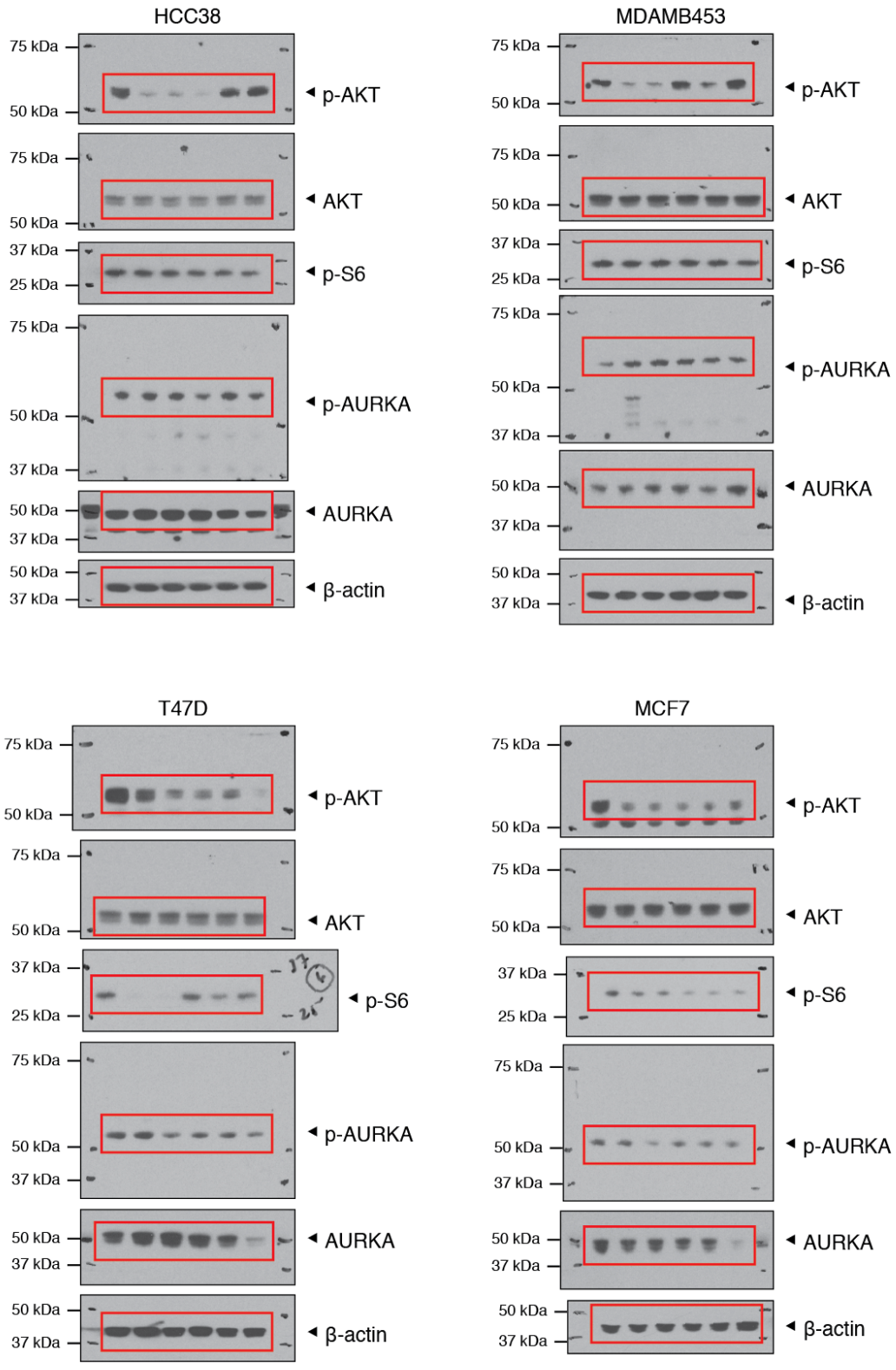



**Supplementary Figure 2.6 Cell line dose response analysis and correlation with sensitivity.** (**a-d**) Cell lines used for the MIBs screening panel were grown in the presence of the indicated compounds and allowed to proliferate for 72 h. Relative proliferation compared to DMSO was scored using cell count and based on $n$ = 4 biologically independent samples. (**e**) Annotation of drug sensitivity using three drug response categories (resistant, moderately sensitive and highly sensitive) based on the indicated $IC_{50}$ values. Because $IC_{50}$ was not reached for PD0325901, $IC_{25}$ is indicated. (**f**) Correlation of kinase activity scores with treatment response. Kinase activity scores were correlated with the following treatment response values: resistant, 3; moderate, 2; sensitive, 1. Positive correlations indicate kinases whose activity was higher in resistant treatment response conditions. Dotted line indicates $P$ = 0.05 based on the $P$ value of correlation. Each point correlation computed over $n$ = 20 biologically independent samples. (**g**) Activities of AURKA across the three treatment response categories. Each point reflects a single cell line and drug treatment, for a total of $n$ = 20 biologically independent samples. Box-and-whisker plots show median (centerline), upper/lower quartiles (box limits), and whiskers spanning the interquartile range from 25-75 percentiles. $P$ values calculated using a two-sided $t$-test.

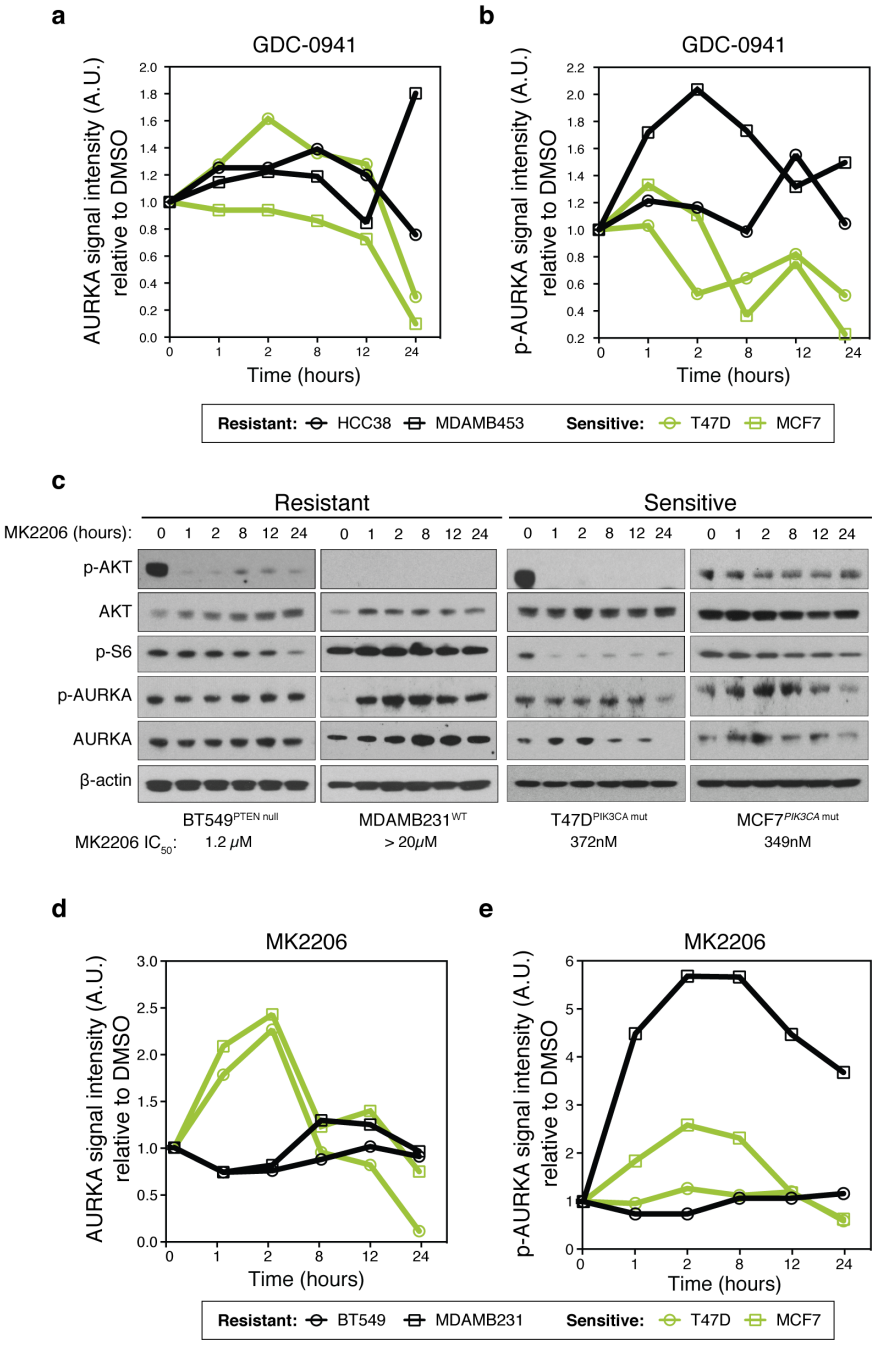



**Supplementary Figure 2.7 Drug-specific analysis of MIBs/MS data.** (a) Ranked lists of kinome dynamics comparing drug sensitive cells versus resistant cells for each drug independently. For

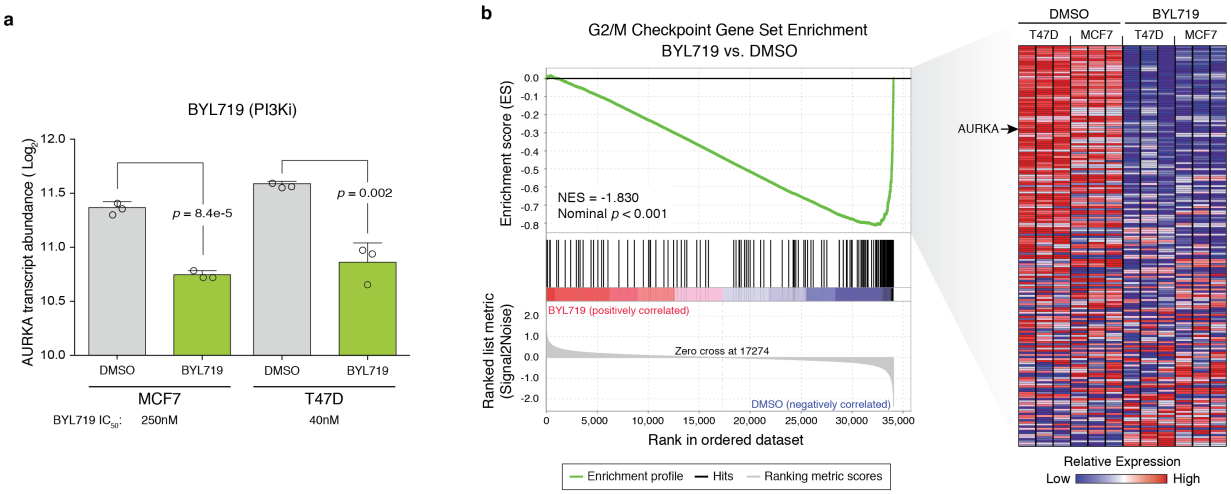
each kinase the level of change was expressed as log ratio of treated/DMSO and this ratio compared between sensitive and resistant cell lines. Kinases were scored independently for each drug (drug-specific analysis; $n$ = 98 GDC-0941, $n$ = 98 MK2206, $n$ = 106 Lapatinib, $n$ = 101 PD0325901) or using a pooled approach (drug-agnostic analysis; $n$ = 75 pooled). Significant differences between sensitive and resistant cells are starred ($P < 0.05$). (**b**) Levels of AURKA defined as $\log_2$ fold change (treatment/DMSO) in sensitive (S) or resistant (R) cells to the indicated inhibitors, displayed by individual drugs or pooled together among all drugs surveyed. Each point reflects a single cell line and drug treatment, for a total of $n$ = 20 biologically independent samples pooled together. Bars represent data mean and $P$ values calculated using a two-sided $t$-test.



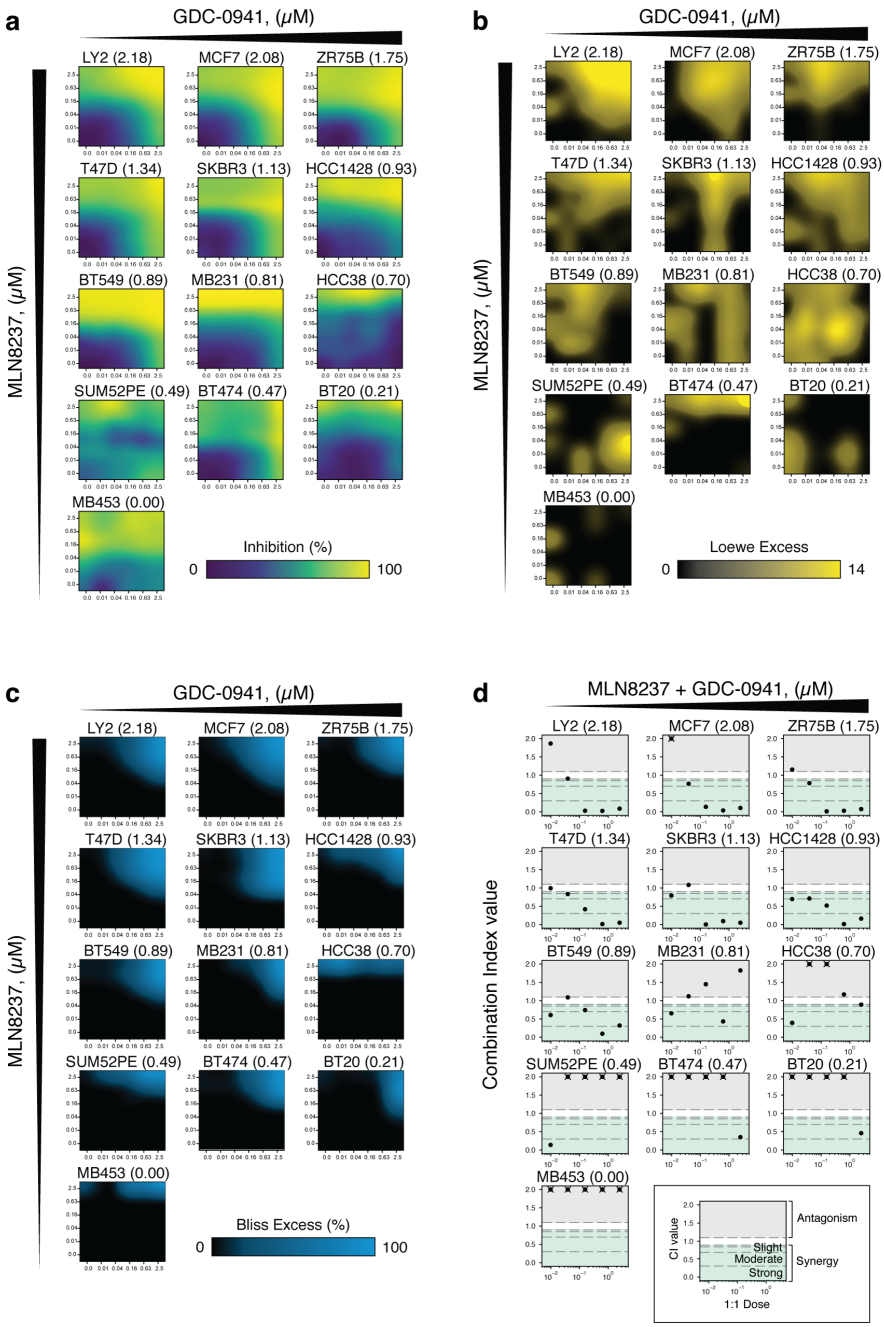

**Supplementary Figure 2.8 Full blot images of Figure 2.2b.** Regions surrounded by red boxes are presented in Fig. 2.2b of the main text.



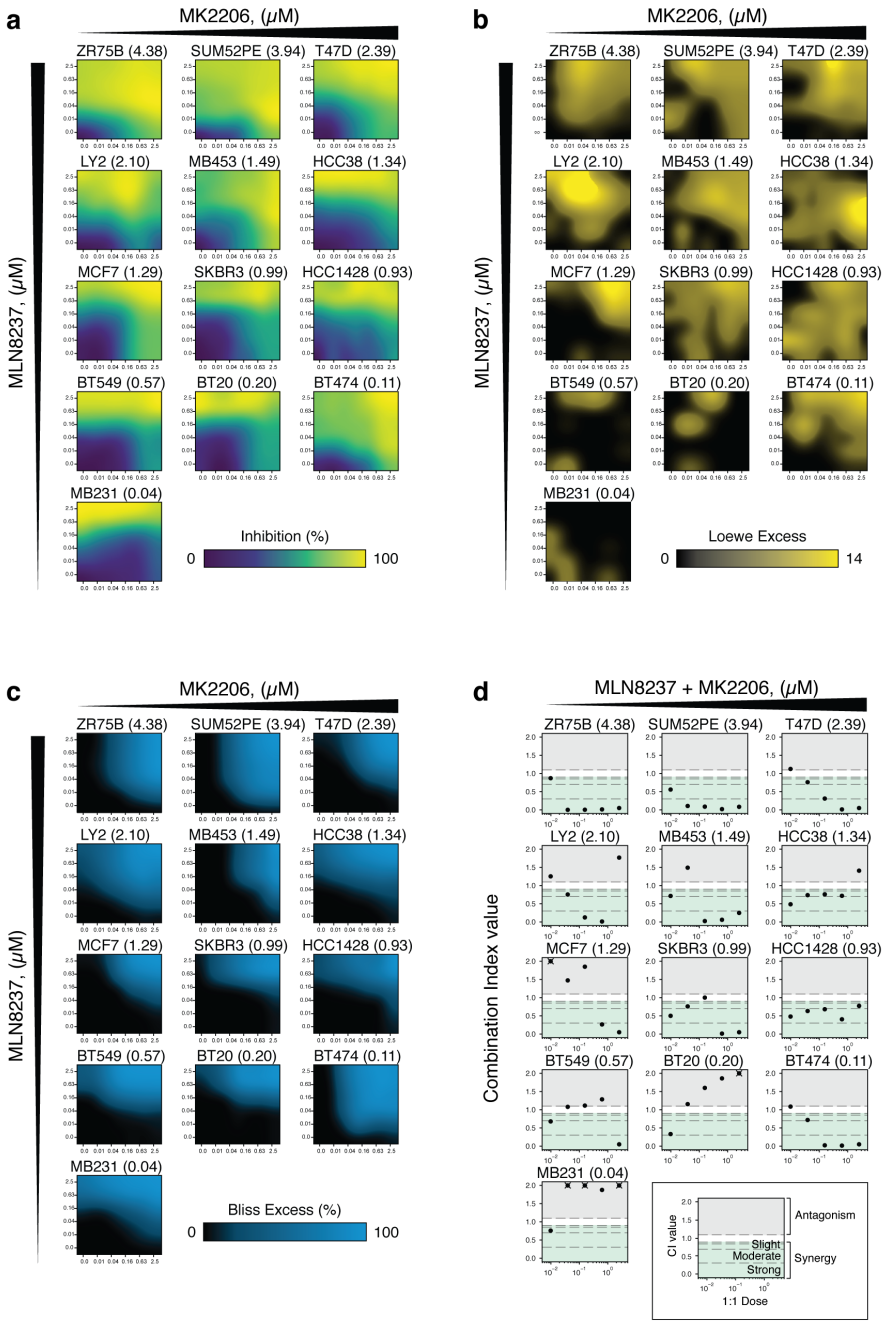

**Supplementary Figure 2.9 GDC-0941 and MK2206 validation in sensitive and resistant cell lines.** (**a**,**b**) Quantification from Fig. 2.2b of total AURKA (**a**) and p-AURKA (**b**) levels from cells treated with 1 μM GDC-0941 over 24 h. (**c**) Immunoblot analysis of PI3K-pathway and AURKA signaling from protein lysates of cell lines treated with 1 μM MK2206 for the indicated periods of time. (**d**,**e**) Quantification of total AURKA (**d**) and p-AURKA (**e**) levels of cell lines treated with 1 μM MK2206 over 24 h from panel (**c**). For all quantifications, densitometry was performed using ImageJ software, and signal intensities were normalized to β-actin loading controls. (**a**-**e**) are representative of $n$ = 3 independent experiments with similar results.



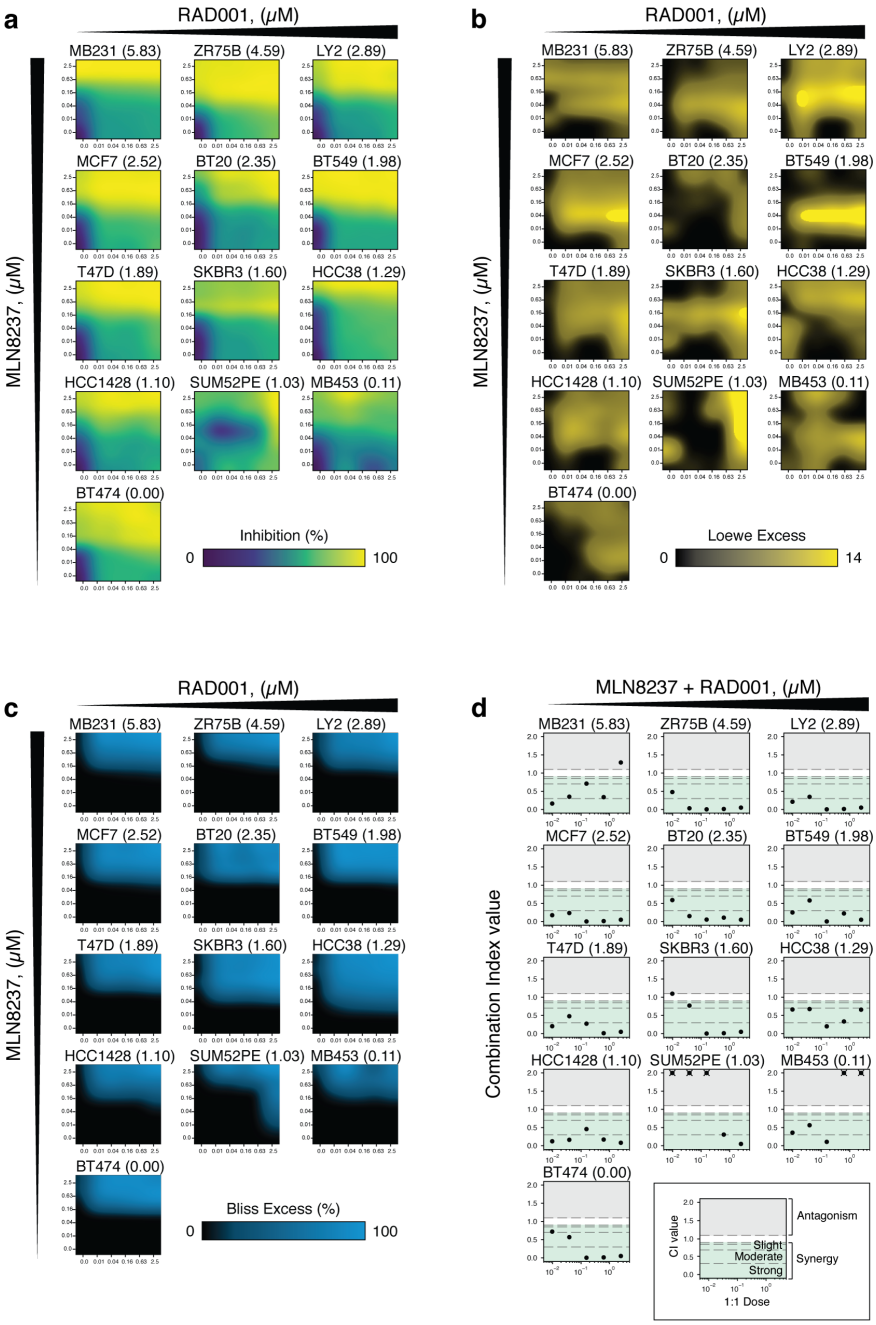

**Supplementary Figure 2.10 PI3K inhibition suppresses transcription of factors involved in G2/M checkpoint.** (**a**) Relative expression of AURKA mRNA from MCF7 or T47D cells treated with 1 μM BYL719 for 24 h ($n$ = 3 biologically independent samples). $IC_{50}$ from Leroy, et al.[18] Error bars are mean ± s.d. and $P$ values calculated using a two-sided $t$-test. (**b**) Leading edge analysis of expression of genes involved in the G2/M checkpoint gene signature. Negative normalized enrichment score (NES) indicates that genes in this set were enriched to be downregulated in response to BYL719 treatment. Normalized transcript levels for 190 genes in this set measured in triplicate is shown. In all graphs, data based on transcriptomic data from Bosch, et al.[17]



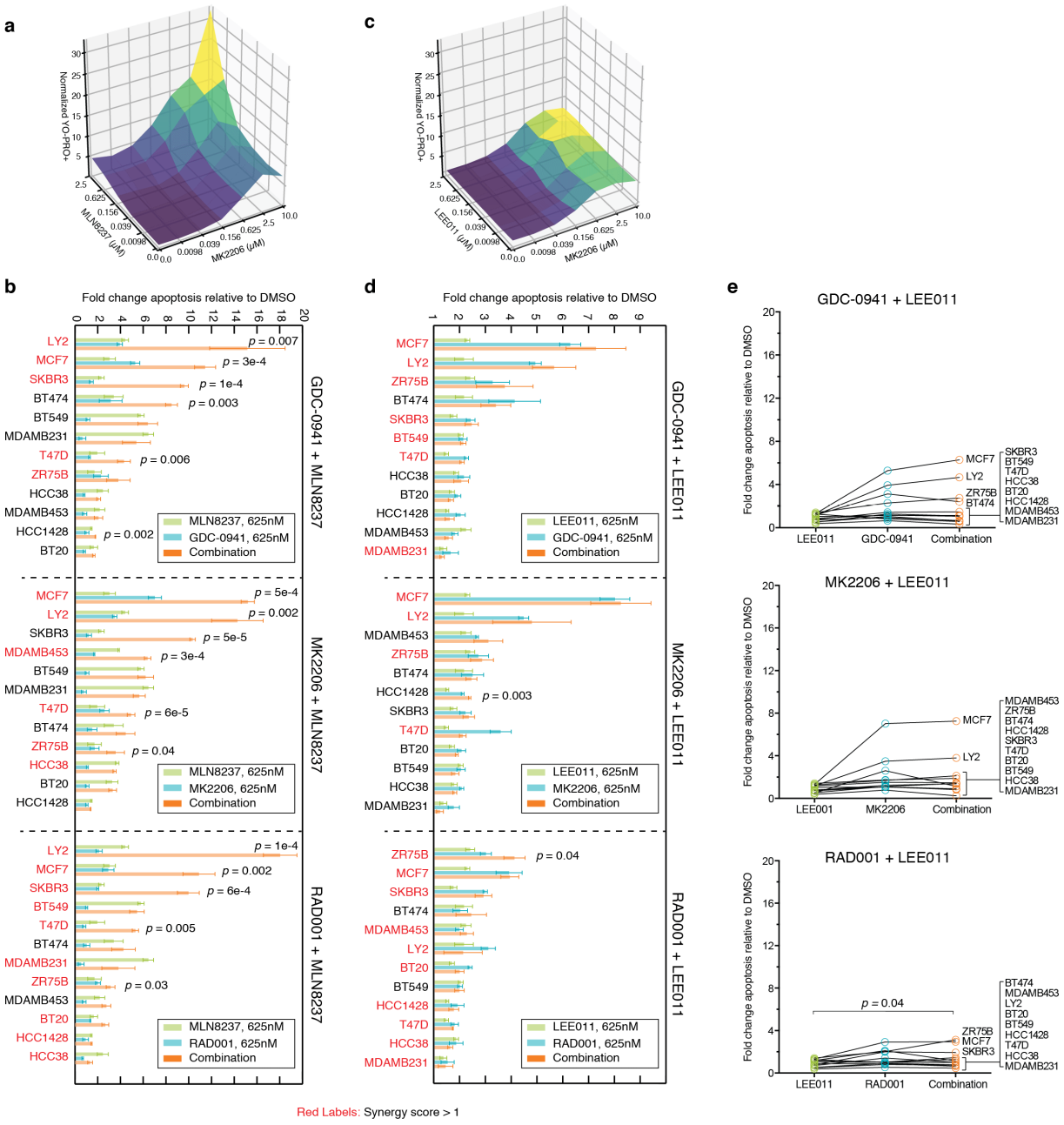

**Supplementary Figure 2.11 Drug synergy analysis for combinations of GDC-0941 and the Aurora kinase inhibitor MLN8237.** Analysis of proliferation data from cells treated with the indicated drug combinations for 72 h. (**a**) Dose-escalation matrices showing percent growth inhibition of cell lines treated with MLN8237 in combination with GDC-0941. (**b**) Loewe excess values representing growth inhibition over an additive model. (**c**) Bliss excess values representing growth inhibition assuming drug independence. (**d**) Combination Index (CI) values based on 1:1 dose combinations. Teal regions represent synergistic CI values. Dotted lines indicate areas of strong synergy (CI < 0.3), moderate synergy (0.3-0.7), slight synergy (0.7-1) and antagonism (CI > 1.2). Cell line names are listed along with the derived synergy score in parenthesis. For clarity, exceptionally high CI values were cutoff at 2.




**Supplementary Figure 2.12 Drug synergy analysis for combinations of MK2206 and the Aurora kinase inhibitor MLN8237.** Analysis of proliferation data from cells treated with the indicated drug combinations for 72 h. (**a**) Dose-escalation matrices showing percent growth inhibition of cell lines treated with MLN8237 in combination with MK2206. (**b**) Loewe excess values representing growth inhibition over an additive model. (**c**) Bliss excess values representing growth inhibition assuming drug independence. (**d**) Combination Index (CI) values based on 1:1 dose combinations. Teal regions represent synergistic CI values. Dotted lines indicate areas of strong synergy (CI < 0.3), moderate synergy (0.3-0.7), slight synergy (0.7-1) and antagonism (CI > 1.2). Cell line names are listed along with the derived synergy score in parenthesis. For clarity, exceptionally high CI values were cutoff at 2.

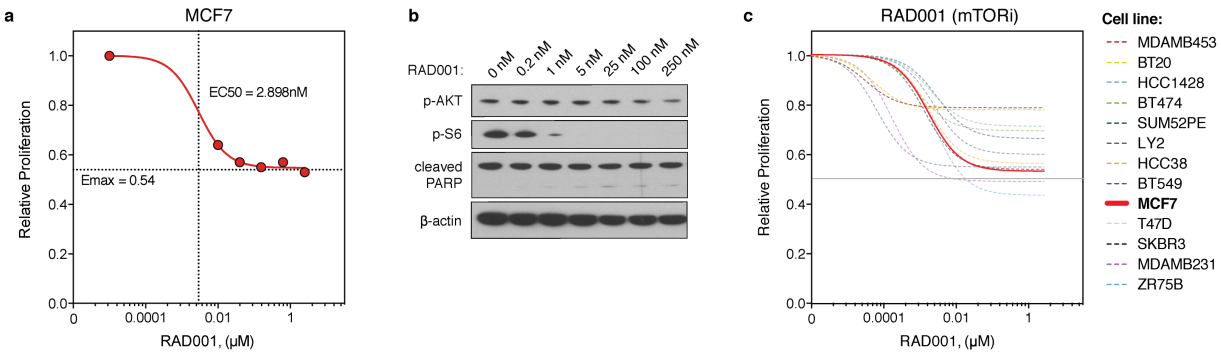



**Supplementary Figure 2.13 Drug synergy analysis for combinations of RAD001 and the Aurora kinase inhibitor MLN8237.** Analysis of proliferation data from cells treated with the indicated drug combinations for 72 h. (**a**) Dose-escalation matrices showing percent growth inhibition of cell lines treated with MLN8237 in combination with RAD001. (**b**) Loewe excess values representing growth inhibition over an additive model. (**c**) Bliss excess values representing growth inhibition assuming drug independence. (**d**) Combination Index (CI) values based on 1:1 dose combinations. Teal regions represent synergistic CI values. Dotted lines indicate areas of strong synergy (CI < 0.3), moderate synergy (0.3-0.7), slight synergy (0.7-1) and antagonism (CI > 1.2). Cell line names are listed along with the derived synergy score in parenthesis. For clarity, exceptionally high CI values were cutoff at 2.

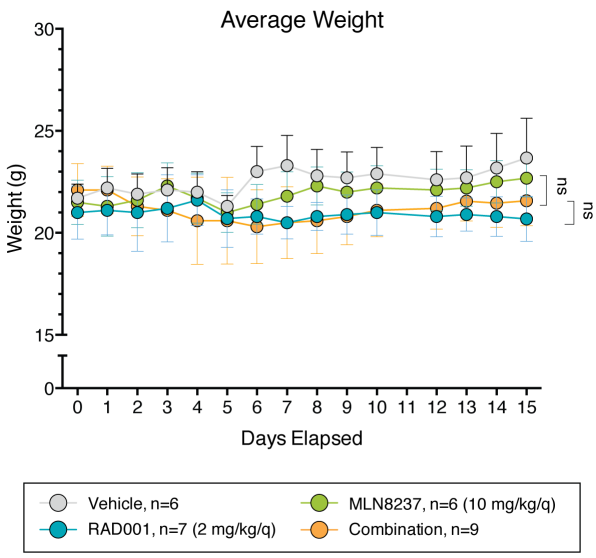



**Supplementary Figure 2.14 Induction of apoptosis in selected single agent and drug combinations.** (**a**) Surface plot of YO-PRO1 positivity for dose-escalation matrix of the combination of MK2206 and MLN8237 in MCF7 cells treated with the indicated dose for 72 h. (**b**) Fold change in YO-PRO1 positivity compared to DMSO for indicated single agent and 1:1 combinations. *P* values shown for cases where the level of apoptosis in the combination was significantly higher than either single agent alone. Red labels indicate cell lines which were deemed to be synergistic based on a synergy score > 1. (**c**) Surface plot of YO-PRO1 positivity across the dose-escalation matrix and (**d**) fold change in apoptosis for cells treated with indicated single agent and combinations with the CDK4/6 inhibitor LEE011. (**e**) Apoptosis compared to

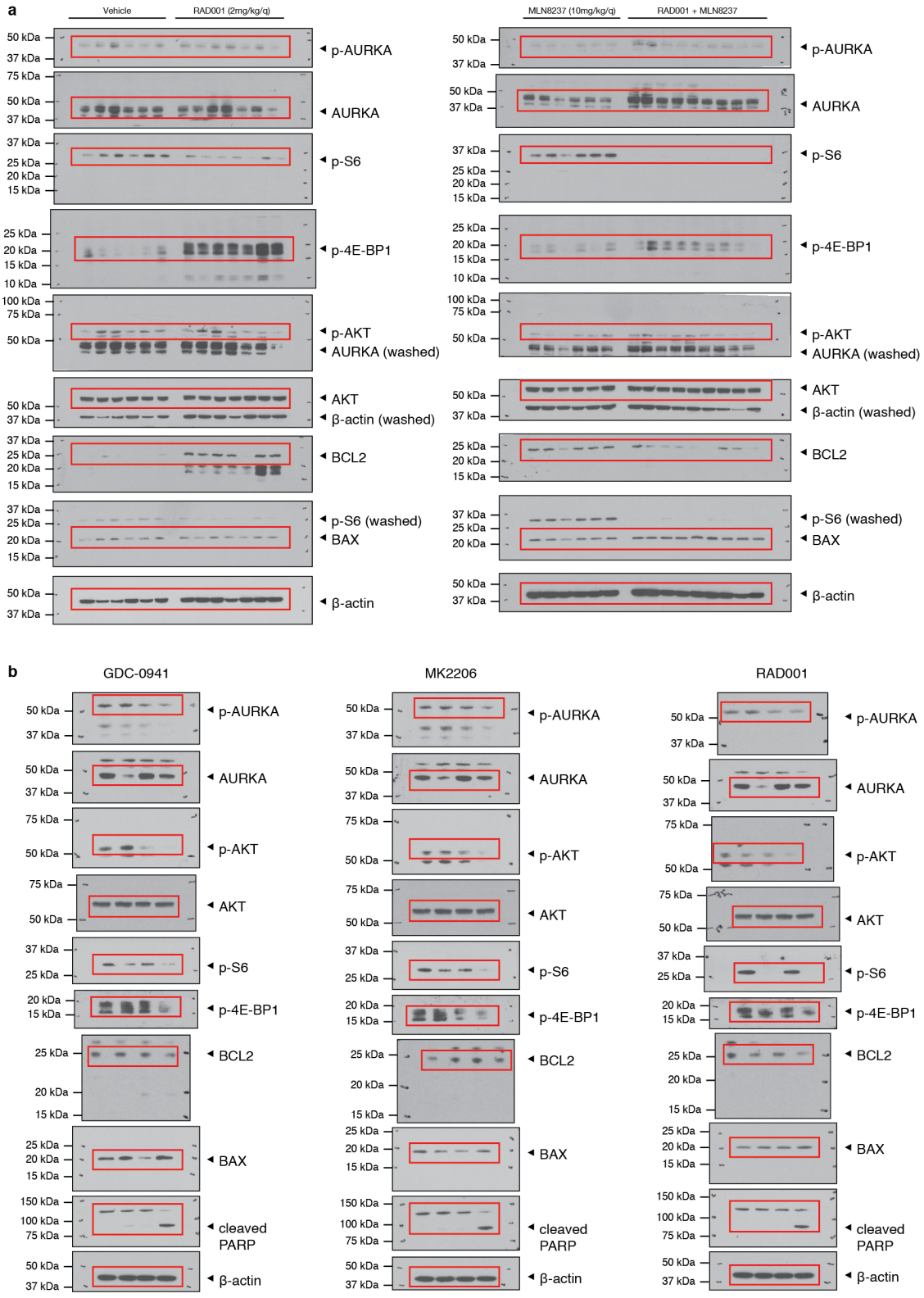
DMSO treatment in cell lines treated with 625 nM of the indicated single agents or combined for 72 h. Apoptosis measured by YO-PRO1 positivity. In all graphs, data are based on $n = 4$ biologically independent samples, error bars are mean ± s.d. and $P$ values calculated using a two-sided $t$-test.



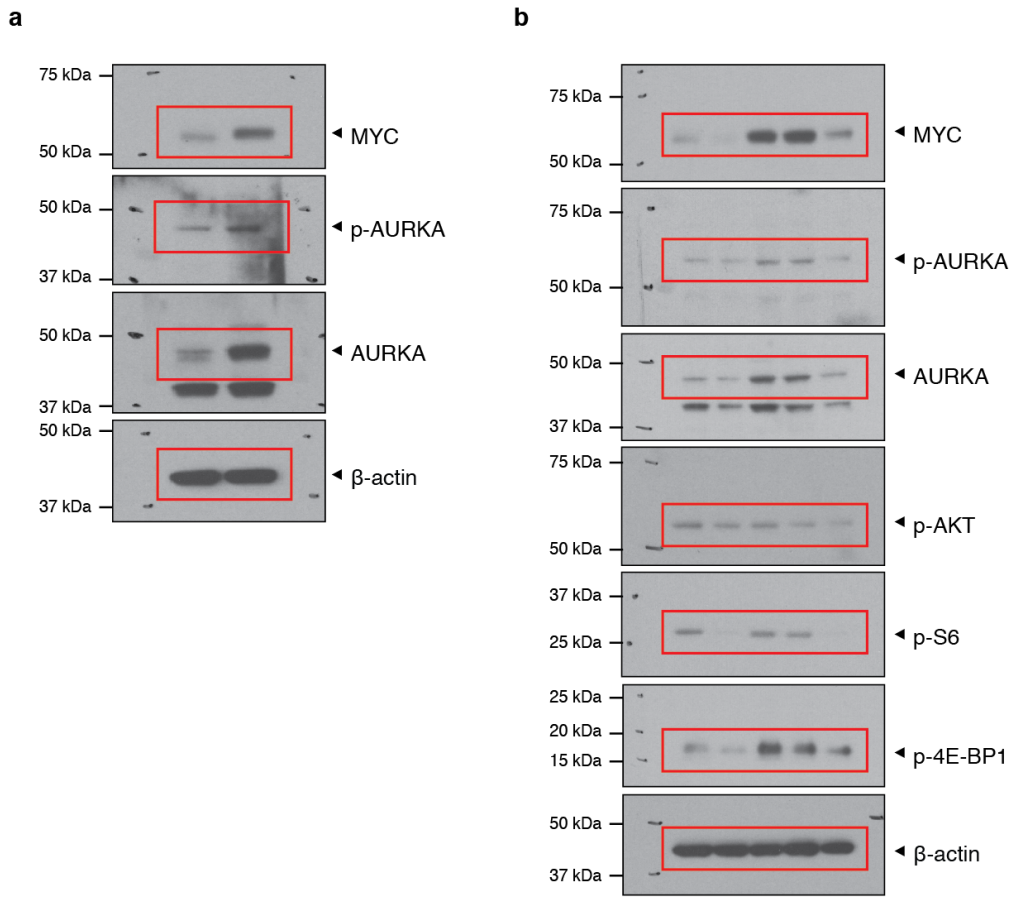

**Supplementary Figure 2.15 RAD001 is cytostatic in MCF7 cells.** (**a**) Dose response analysis of MCF7 cells exposed to RAD001 for 72 h. Proliferation measured by cell count from $n$ = 4 biologically independent samples and error bars are mean ± s.d. (**b**) MCF7 cells were treated with the indicated doses of RAD001 for 24 h. Protein lysates were subjected to immunoblot and analyzed for PI3K-pathway signaling and cell death using the indicated antibodies. Representative image of $n$ = 2 independent experiments with similar results. (**c**) Dose response analysis of proliferation for a panel of 13 breast cancer cell lines following 72 h exposure to RAD001 from data based on $n$ = 4 biologically independent samples.



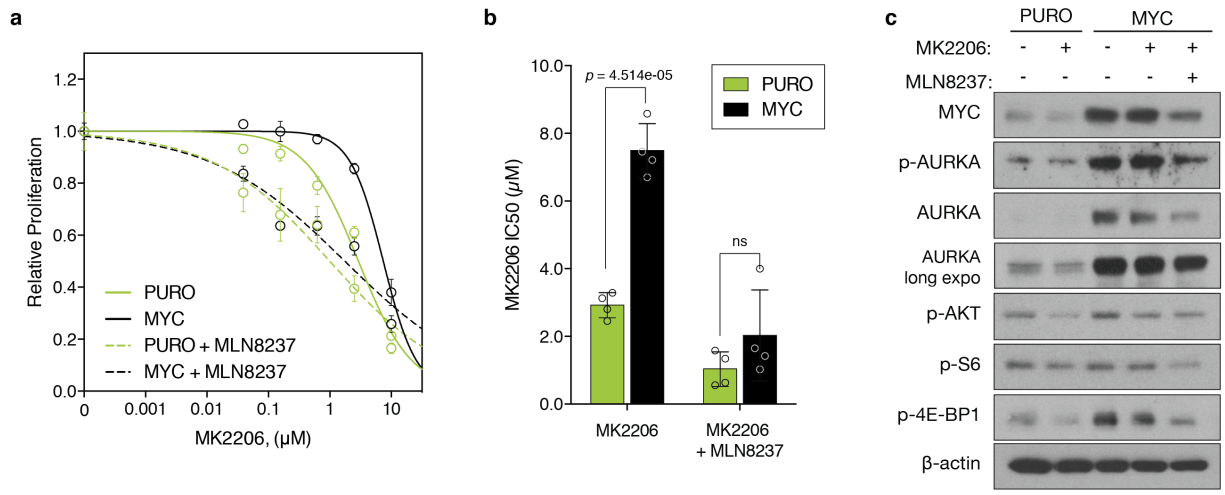

**Supplementary Figure 2.16 Mouse weight during xenograft studies.** Average mouse weight over the course of 15 d treatment with vehicle ($n$ = 6 biologically independent mice) or the indicated dose of RAD001 ($n$ = 7 biologically independent mice), MLN8237 ($n$ = 6 biologically independent mice) or the combination ($n$ = 9 biologically independent mice). Error bars are mean ± s.d. and $P$ values calculated using a two-sided $t$-test.




**Supplementary Figure 2.17 Full blot images of Figure 2.5.** Regions surrounded by red boxes are presented in (**a**) Fig. 2.5a and (**b**) Fig. 2.5b of the main text.




**Supplementary Figure 2.18 Full blot images of Figure 2.6.** Regions surrounded by red boxes are presented in (**a**) Fig. 2.6d and (**b**) Fig. 2.6g of the main text.




**Supplementary Figure 2.19 MYC expression causes resistance to AKT inhibition via AURKA mediated activation of the PI3K pathway.** (**a**) Proliferation of control (PURO) or MYC MCF10A cells in response to MK2206 or treated with the combination of 2.5 nM MLN8237. Combinations were normalized to MLN8237 alone. Data represents $n$ = 4 biologically independent samples. (**b**) $IC_{50}$ analysis of dose-response curves shown in (**a**) from $n$ = 4 independent samples. (**c**) Immunoblot analysis of lysates from control and MYC MCF10A cells treated for 24 h with 1 μM of MK2206 as a single-agent or in combination with 100 nM of MLN8237. Representative image of $n$ = 3 independent experiments with similar results. In all graphs, error bars are mean ± s.d. and $P$ values were calculated using a two-sided $t$-test.

## 2.10 References


1. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
2. Janku, F. *et al.* Assessing PIK3CA and PTEN in Early-Phase Trials with PI3K/AKT/mTOR Inhibitors. *Cell Rep.* **6**, 377–387 (2014).
3. Dey, N., De, P. & Leyland-Jones, B. PI3K-AKT-mTOR inhibitors in breast cancers: From tumor cell signaling to clinical trials. *Pharmacol. Ther.* **175**, 91–106 (2017).
4. O'Reilly, K. E. *et al.* mTOR Inhibition Induces Upstream Receptor Tyrosine Kinase Signaling and Activates Akt. *Cancer Res.* **66**, 1500–1508 (2006).
5. Klempner, S. J., Myers, A. P. & Cantley, L. C. What a tangled web we weave: emerging resistance mechanisms to inhibition of the phosphoinositide 3-kinase pathway. *Cancer Discov.* **3**, 1345–1354 (2013).
6. Shah, P. D. & Chandarlapaty, S. Resistance to PI3K Pathway Inhibition. in *PI3K-mTOR in Cancer and Cancer Therapy* (eds. Dey, N., De, P. & Leyland-Jones, B.) 125–147 (Springer International Publishing, 2016). doi:10.1007/978-3-319-34211-5_5.
7. Duncan, J. S. *et al.* Dynamic Reprogramming of the Kinome in Response to Targeted MEK Inhibition in Triple-Negative Breast Cancer. *Cell* **149**, 307–321 (2012).
8. Corcoran, R. B. *et al.* TORC1 Suppression Predicts Responsiveness to RAF and MEK Inhibition in BRAF-Mutant Melanoma. *Sci. Transl. Med.* **5**, 196ra98-196ra98 (2013).

9. Anderson, G. R. *et al.* PIK3CA mutations enable targeting of a breast tumor dependency through mTOR-mediated MCL-1 translation. *Sci. Transl. Med.* **8**, 369ra175-369ra175 (2016).

10. Boussemart, L. *et al.* eIF4F is a nexus of resistance to anti-BRAF and anti-MEK cancer therapies. *Nature* **513**, 105–109 (2014).

11. Elkabets, M. *et al.* mTORC1 Inhibition Is Required for Sensitivity to PI3K p110α Inhibitors in PIK3CA-Mutant Breast Cancer. *Sci. Transl. Med.* **5**, 196ra99-196ra99 (2013).

12. Vora, S. R. *et al.* CDK 4/6 Inhibitors Sensitize PIK3CA Mutant Breast Cancer to PI3K Inhibitors. *Cancer Cell* **26**, 136–149 (2014).

13. Bantscheff, M. *et al.* Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat. Biotechnol.* **25**, 1035–1044 (2007).

14. Sos, M. L. *et al.* Oncogene Mimicry as a Mechanism of Primary Resistance to BRAF Inhibitors. *Cell Rep.* **8**, 1037–1048 (2014).

15. Stuhlmiller, T. J. *et al.* Inhibition of Lapatinib-Induced Kinome Reprogramming in ERBB2-Positive Breast Cancer by Targeting BET Family Bromodomains. *Cell Rep.* **11**, 390–404 (2015).

16. Lens, S. M. A., Voest, E. E. & Medema, R. H. Shared and separate functions of polo-like kinases and aurora kinases in cancer. *Nat. Rev. Cancer* **10**, 825–841 (2010).

17. Bosch, A. *et al.* PI3K inhibition results in enhanced estrogen receptor function and dependence in hormone receptor–positive breast cancer. *Sci. Transl. Med.* **7**, 283ra51-283ra51 (2015).

18. Leroy, C. *et al.* Activation of IGF1R/p110β/AKT/mTOR confers resistance to α-specific PI3K inhibition. *Breast Cancer Res. BCR* **18**, 41 (2016).

19. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).

20. Foucquier, J. & Guedj, M. Analysis of drug combinations: current methodological landscape. *Pharmacol. Res. Perspect.* **3**, e00149 (2015).

21. Fruman, D. A. & Rommel, C. PI3K and cancer: lessons, challenges and opportunities. *Nat. Rev. Drug Discov.* **13**, 140–156 (2014).

22. Nikonova, A. S., Astsaturov, I., Serebriiskii, I. G., Dunbrack, R. L. & Golemis, E. A. Aurora-A kinase (AURKA) in normal and pathological cell growth. *Cell. Mol. Life Sci. CMLS* **70**, 661–687 (2013).

23. Piccart, M. *et al.* Everolimus plus exemestane for hormone-receptor-positive, human epidermal growth factor receptor-2-negative advanced breast cancer: overall survival results from BOLERO-2†. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **25**, 2357–2362 (2014).

24. Choo, A. Y., Yoon, S.-O., Kim, S. G., Roux, P. P. & Blenis, J. Rapamycin differentially inhibits S6Ks and 4E-BP1 to mediate cell-type-specific repression of mRNA translation. *Proc. Natl. Acad. Sci.* **105**, 17414–17419 (2008).

25. Vadlakonda, L., Dash, A., Pasupuleti, M., Anil Kumar, K. & Reddanna, P. The Paradox of Akt-mTOR Interactions. *Front. Oncol.* **3**, 165 (2013).

26. Laplante, M. & Sabatini, D. M. mTOR Signaling in Growth Control and Disease. *Cell* **149**, 274–293 (2012).

27. Volkmann, N., Marassi, F. M., Newmeyer, D. D. & Hanein, D. The rheostat in the membrane: BCL-2 family proteins and apoptosis. *Cell Death Differ.* **21**, 206–215 (2014).

28. West, M. J., Stoneley, M. & Willis, A. E. Translational induction of the c-***myc*** oncogene via activation of the FRAP/TOR signalling pathway. *Oncogene* **17**, 769 (1998).

29. Sears, R. *et al.* Multiple Ras-dependent phosphorylation pathways regulate Myc protein stability. *Genes Dev.* **14**, 2501–2514 (2000).

30. Diehl, J. A., Cheng, M., Roussel, M. F. & Sherr, C. J. Glycogen synthase kinase-3beta regulates cyclin D1 proteolysis and subcellular localization. *Genes Dev.* **12**, 3499–3511 (1998).

31. Yang, H., He, L., Kruk, P., Nicosia, S. V. & Cheng, J. Q. Aurora-A induces cell survival and chemoresistance by activation of Akt through a p53-dependent manner in ovarian cancer cells. *Int. J. Cancer* **119**, 2304–2312 (2006).

32. Dar, A. A., Belkhiri, A. & El-Rifai, W. The aurora kinase A regulates GSK-3beta in gastric cancer cells. *Oncogene* **28**, 866–875 (2009).

33. Martins, M. M. *et al.* Linking Tumor Mutations to Drug Responses via a Quantitative Chemical–Genetic Interaction Map. *Cancer Discov.* **5**, 154–167 (2015).

34. Muellner, M. K. *et al.* A chemical-genetic screen reveals a mechanism of resistance to PI3K inhibitors in cancer. *Nat. Chem. Biol.* **7**, 787 (2011).

35. Ilic, N., Utermark, T., Widlund, H. R. & Roberts, T. M. PI3K-targeted therapy can be evaded by gene amplification along the MYC-eukaryotic translation initiation factor 4E (eIF4E) axis. *Proc. Natl. Acad. Sci.* **108**, E699–E708 (2011).

36. Liu, P. *et al.* Oncogenic PIK3CA-driven mammary tumors frequently recur via PI3K pathway–dependent and PI3K pathway–independent mechanisms. *Nat. Med.* **17**, 1116 (2011).

37. Nomanbhoy, T. K. *et al.* Chemoproteomic Evaluation of Target Engagement by the Cyclin-Dependent Kinase 4 and 6 Inhibitor Palbociclib Correlates with Cancer Cell Response. *Biochemistry* **55**, 5434–5441 (2016).

38. Roux, P. P. *et al.* RAS/ERK signaling promotes site-specific ribosomal protein S6 phosphorylation via RSK and stimulates cap-dependent translation. *J. Biol. Chem.* **282**, 14056–14064 (2007).

39. She, Q.-B. *et al.* 4E-BP1 is a key effector of the oncogenic activation of the AKT and ERK signaling pathways that integrates their function in tumors. *Cancer Cell* **18**, 39–51 (2010).

40. Hollander, J. den *et al.* Aurora kinases A and B are up-regulated by Myc and are essential for maintenance of the malignant state. *Blood* **116**, 1498–1505 (2010).

41. Lu Longfeng *et al.* Aurora kinase A mediates c-Myc's oncogenic effects in hepatocellular carcinoma. *Mol. Carcinog.* **54**, 1467–1479 (2014).

42. Zheng, F. *et al.* Nuclear AURKA acquires kinase-independent transactivating function to enhance breast cancer stem cell phenotype. *Nat. Commun.* **7**, 10180 (2016).

43. Ferrell, J. E. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr. Opin. Cell Biol.* **14**, 140–148 (2002).

44. Katsha, A. *et al.* Activation of EIF4E by Aurora Kinase A Depicts a Novel Druggable Axis in Everolimus-Resistant Cancer Cells. *Clin. Cancer Res.* **23**, 3756–3768 (2017).

45. Savannah, K. J. B. *et al.* Dual Targeting of mTOR and Aurora-A Kinase for the Treatment of Uterine Leiomyosarcoma. *Clin. Cancer Res.* **18**, 4633–4645 (2012).

46. Liu, L.-L. *et al.* Inhibition of mTOR Pathway Sensitizes Acute Myeloid Leukemia Cells to Aurora Inhibitors by Suppression of Glycolytic Metabolism. *Mol. Cancer Res.* **11**, 1326–1336 (2013).

47. Baselga, J. *et al.* Buparlisib plus fulvestrant versus placebo plus fulvestrant in postmenopausal, hormone receptor-positive, HER2-negative, advanced breast cancer

(BELLE-2): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol.* (2017) doi:10.1016/S1470-2045(17)30376-5.

48. Zwang, Y. *et al.* Synergistic interactions with PI3K inhibition that induce apoptosis. *eLife* **6**, e24523 (2017).

49. Barr, P. M. *et al.* Phase II Intergroup Trial of Alisertib in Relapsed and Refractory Peripheral T-Cell Lymphoma and Transformed Mycosis Fungoides: SWOG 1108. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **33**, 2399–2404 (2015).

50. Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**, 515–527 (2006).

51. Barvian, M. *et al.* Pyrido[2,3-d]pyrimidin-7-one Inhibitors of Cyclin-Dependent Kinases. *J. Med. Chem.* **43**, 4606–4616 (2000).

52. Daub, H. *et al.* Kinase-Selective Enrichment Enables Quantitative Phosphoproteomics of the Kinome across the Cell Cycle. *Mol. Cell* **31**, 438–448 (2008).

53. Blake, J. F. *et al.* Discovery of pyrrolopyrimidine inhibitors of Akt. *Bioorg. Med. Chem. Lett.* **20**, 5607–5612 (2010).

54. Tanaka, M. *et al.* An unbiased cell morphology-based screen for new, biologically active small molecules. *PLoS Biol.* **3**, e128 (2005).

55. Bankston, D. *et al.* A Scaleable Synthesis of BAY 43-9006: A Potent Raf Kinase Inhibitor for the Treatment of Cancer. *Org. Process Res. Dev.* **6**, 777–781 (2002).

56. Statsuk, A. V. *et al.* Tuning a Three-Component Reaction For Trapping Kinase Substrate Complexes. *J. Am. Chem. Soc.* **130**, 17568–17574 (2008).

57. MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinforma. Oxf. Engl.* **26**, 966–968 (2010).

58. Choi, M. *et al.* MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* **30**, 2524–2526 (2014).

59. Lehár, J. *et al.* Synergistic drug combinations tend to improve therapeutically relevant selectivity. *Nat. Biotechnol.* **27**, 659–666 (2009).

# Chapter 3.

# Single-cell pharmacogenomics targets an inflamed chemoresistant tumor subpopulation in triple negative breast cancer

## 3.1 Abstract

Cancers often harbor high levels of intra-tumoral heterogeneity, or differences among individual cancer cells, which may directly contribute to drug resistance through the selection of mutant subclones. We hypothesized that non-genetic differences in cellular states between individual cancer cells may also reflect altered drug sensitivities and as such, can be used to inform upfront drug combinations that elicit more complete anti-tumor responses. To capture functional heterogeneity we performed single-cell RNA sequencing (scRNAseq) of 5 different *in vitro* and *in vivo* breast cancer models encompassing 55,000 breast cancer cells. We identified tumor cell subpopulations reflective of different subtypes or cellular states associated with resistance to standard-of-care therapies in all models analyzed. Notably, we identified a recurrent subpopulation present in all triple-negative breast cancer (TNBC) models where a rare fraction of cells exist in a reversible interferon-stimulated gene (ISG)-state caused by cGAS/STING signaling. Pharmacogenomic modeling based on bulk cancer cell line drug sensitivities predicted that ISG cells were chemotherapy resistant and hypersensitive to CDK4/6 and mTOR

inhibition, which we experimentally validated. ISG expression is associated with poor prognosis and residual disease after chemotherapy in TNBC, implicating this population in disease progression and highlighting the need for identifying, modeling and targeting cancer cells in a pre-existing drug resistant state.

## 3.2 Introduction

Drug resistance is a major problem in breast cancer where insensitivity to chemotherapy is associated with a 40 to 80% risk of recurrence after neo-adjuvant therapy resulting in incurable metastatic disease and death for most patients[1–3]. The need for new therapeutic approaches is particularly pressing in the context of triple-negative breast cancers (TNBC), the most aggressive subtype of disease accounting for 12 to 18% of all breast cancer cases[4], where responses to chemotherapy are not durable. In the case of neo-adjuvant chemotherapy, 30 to 50% of patients with localized TNBC evolve resistance leading to poor overall survival for this group of patients[4,5]. Similar phenomena has also been observed in respect to targeted therapies, as seen with HER2 inhibitor treatment leading to a pathologic complete response (pCR; defined as no invasive or in situ residual tumors in the breast) rate of only 25 to 30% in HER2+ breast cancers[6,7]. Although one mode of drug resistance can be attributed to the Darwinian selection of pre-existing genetically diverse subclones[8–10], clinical evidence from breast cancers undergoing neo-adjuvant chemotherapy have indicated treatment often does not result in subclonal selection demonstrating that it is not necessarily genetically encoded[9,11–13]. These properties are also

evident *in vitro* in breast cancer cell lines (BCCLs) and patient-derived xenograft (PDX) models that can also evolve drug resistance through genetic and non-genetic mechanisms[14–16]. To this end, cells existing in different cell states can possess differential drug sensitivities further complicating the therapeutic management of breast tumors due to the inherent plasticity of tumor cell states[16–19]. Hence new approaches are needed to counteract both genetic and non-genetic forms of drug resistance in breast cancer.

Molecular variability between individual cancer patients (inter-tumoral heterogeneity) can predict who will respond to therapy, such as differences in molecular subtypes, which is crucial in determining predicted responsiveness to standard-of-care therapies[20–22]. As such, classification of intrinsic molecular subtype from bulk patient tumor samples has become an important diagnostic tool to predict risk of relapse and response to chemotherapy in breast cancers. However, patients do not present with homogenous tumors clinically, often times lacking uniform expression of therapeutic targets[23]. This heterogeneity can severely impact the clinical assessment of a tumor and subsequent therapeutic regimens, and may facilitate recurrence whereby subpopulations of cells are spared during treatment. Similarly, molecular variability among individual cancer cells (intra-tumoral heterogeneity) determines which are selected for by therapy. While this variability may be genetic in the form of multiple different subclones within a tumor mass being linked with resistance[9,11–13], this variability may also be non-genetic and even stochastic in nature where breast cancer cells in a drug resistant state are present prior to therapy[9].

Single-cell DNA- and RNA-sequencing methods have emerged as powerful tools for resolving intra-tumoral heterogeneity, reconstructing tumor evolutionary lineages, and characterizing

distinct tumor cell subpopulations in breast cancers[9,24,25]. However, approaches to use these data derived from single-cell studies towards predicting and testing new therapeutic strategies have remained a challenge. Studies linking genomics to drug efficacy, or pharmacogenomics, have relied on the analysis of large panels of cell lines or PDX systems having been both molecularly profiled and functionally challenged through high-throughput drug screens. However, these bulk assays have limited application to studying the functional consequences of cell-cell variability in that heterogeneous responses of cells from within a bulk population and subtle changes in cell state or gene expression are obscured in bulk analysis. Therefore, the functional implications that individual cells within a population pose on therapeutic responses remains unknown due to the lack in feasibility of directly challenging a single live cell with drug. As such, approaches for characterizing genetic and functional variability contributing to drug response at the cellular level are needed to inform upfront drug combinations that elicit more complete anti-tumor responses.

Here we hypothesized that the molecular variability between individual breast cancer cells and subsequent impact on drug sensitivities could be modeled by the differences observed between individual cancer samples and therapeutic outcomes. To test this hypothesis we measured transcript levels in single cells from various breast cancer models both *in vitro* using cell lines and *in vivo* using patient-derived xenograft samples and patient data and identified distinct subpopulations in each model tested that predicted resistance to standard-of-care breast cancer therapies. One of these populations was identified solely in TNBC samples and was comprised of cells in an interferon-stimulated gene (ISG) state whose representative genes were elevated in TNBC patient samples and associated with poor outcomes. We developed a computational approach to identify signatures of drug resistance and sensitivity based on cell line pharmacologic

profiling datasets and patient-derived gene modules, which predicted that ISG cells were resistant to chemotherapy, which was experimentally validated in TNBC cells using a single-live-cell tracking workflow. Pharmacogenomic modeling based on bulk cancer cell line drug sensitivities also predicted that ISG cells were hypersensitive to CDK4/6 and mTOR inhibition, which we confirmed as an effective strategy for targeting this chemoresistant subpopulation. Here we provide an effective new framework to systematically discover drug combinations capable of counteracting tumor heterogeneity by predicting drug responses of individual breast cancer cells. Our study potentially explains the cellular source of heightened ISG expression found in residual TNBC tumors after chemotherapy and systematically links tumor transcriptional heterogeneity with drug actionability to optimize therapy combinations.

## 3.3 Results

### 3.3.1 Single-cell heterogeneity is pervasive across various models of breast cancer

To first understand the functional consequences of heterogeneity on drug resistance, we analyzed the transcript profiles of individual cells from various models of breast cancer including pre-clinical and patient sample data (Fig. 3.1a). Transcriptional profiling was performed at the single-cell level on a panel of cell lines representing various subtypes of breast cancer ($MCF7^{ER+}$, $SKBR3^{HER2}$, MDA-MB-$231^{TNBC}$, $HCC38^{TNBC}$) in addition to freshly dissociated cells from a patient-derived xenograft (PDX) model of TNBC ($HCI\text{-}002^{PDX\text{-}TNBC}$) using the 10x Genomics Chromium™ platform. At an average depth of 23,573 reads per cell, sequencing resulted in an

average number of 6,491 unique molecular identifiers (UMIs) and 1,609 genes detected per cell following initial quality control filters (Supplementary Dataset 3.1). Overall, we studied the expression profiles of 54,599 single cells (3,080 – 19,173 cells per sample) from 5 breast cancer models reflecting the different subtypes of disease. Additionally, we obtained single-cell profiles of longitudinal tumor samples from a cohort of 8 TNBC patients[9] both pre- and-post chemotherapy treatment. Together this data constitutes an extensive representation of breast cancer heterogeneity and provides a useful resource for studying its functional implications.

To investigate the magnitude of heterogeneity across breast cancer, we performed an unsupervised graph-based clustering of all cells pooled from the panel of breast cancer cell lines ($n$ = 51,519 cells) using genes from significant principal components, and visualizing the space using Uniform Manifold Approximation and Projection (UMAP)[26] (Fig. 3.1b). To avoid batch effects, all samples were processed together with the exception of HCC38 data. Comparison of single cell profiles using a two dimensional UMAP embedding revealed significant heterogeneity in cells to the extent that often times cells from one model had high similarity with cells from a different model, even through different models were derived from different patients with distinct disease subtypes (Fig. 3.1b). To this end, despite cancer cell lines being accepted homogenous model systems, we observed a number of cells from every cell line clustering apart from their predominant parent cluster. For example, cells from the luminal breast cancer cell line $MCF7^{ER+}$ were found in clusters primarily composed of cells from a basal tumor model. TNBC MDA-MB-231 cells also appeared highly heterogeneous in that a subpopulation of cells were found clustered with both ER+ and HER2-amplified dominated subclusters (Fig. 3.1b). These results were largely robust with respect to various methods for measuring cell-cell similarity and

clustering (Supplementary Fig. 3.1). Taken together, based on the transcript profiles of individual cells this data suggests the existence of subpopulations within homogenous models of breast cancer that identify with and may functionally mimic cell types of a different lineage separate from the bulk tumor classification. This finding may have significant clinical implications given that breast cancer standard-of-care treatment regimens are typically prescribed based on subtype classifications. We next investigated if this degree of heterogeneity is reflected across patient tumor cells by looking within single-cell profiles from a cohort of TNBC patients[9]. Here we also observed similarities in cell states between 2,640 single cells derived from 8 treatment-naïve TNBC patients, indicating that the degree of between tumor (inter-tumoral) heterogeneity is at least partly matched by the heterogeneity between cancer cells from the same tumor (intra-tumoral heterogeneity). Rather than cells from each individual patient clustering based on parent sample identity, we observed significant overlap between groups where cells from various patients were intermixed in different subclusters ($n$ = 2,640 pre-treatment cells; Fig. 3.1c). Previous work analyzing the gene expression profiles of 515 breast tumor cells from 11 patients also found extensive intra-tumoral heterogeneity, corroborating a mixed representation of intra- and inter-patient cells[24]. This suggests that transcriptional diversity occurs in all levels of breast cancer models spanning cell lines to patient profiles.

We next asked if this degree of heterogeneity is observed even deeper within individual cell lines and samples. To identify distinct subpopulations present within specific samples, we performed clustering analysis based on the variation in gene expression across cells within individual cell lines and PDX models. Following the Seurat pipeline[27,28], we applied a principal component analysis on the top significant variable genes ranked by their normalized dispersion and

performed optimized *K*-means clustering on the PCs to categorize captured cells into distinct subpopulations, as described previously[29]. The identified subclusters were then visualized using t-Distributed Stochastic Neighbor Embedding (tSNE)[30] with cells partitioning into 6 – 9 subpopulations within each model (Fig. 3.1d,e; Supplementary Dataset 3.2). Notably, we observed co-existing subpopulations present within each sample tested revealing widespread heterogeneity at the transcript level in breast cancer. We then inferred copy number variations (CNV)[31] called using these data to determine the degree to which transcriptionally distinct subpopulations could be attributed to changes in clonality as measured by CNV levels (Supplementary Fig. 3.2). Ongoing analyses show multiple genetic subclones are present within subpopulations identified using expression-based clustering and hence, some but not all variability between cancer cells can be attributed to subclonality.

### 3.3.2 Modular analysis of individual cells informs functional heterogeneity and mirrors inter-patient variability

Having observed extensive heterogeneity at the single-cell level across the breast cancer models we profiled, we next sought to characterize this variability in terms of molecular mechanisms and function in addition to impacts on therapeutic response. To do so, we developed a new approach using a Modular Drug Response Prediction Pipeline ($MOD_{Rx}$) in order to systematically discover drug combinations capable of counteracting tumor heterogeneity by predicting drug responses of individual breast cancer cells (Fig. 3.2a). We began by performing an in-depth molecular characterization of the luminal ER+ MCF7 cell line to study the extent that various subpopulations exist in different functional states. An unbiased analysis of gene expression

profiles of individual MCF7$^{ER+}$ cells identified 7 subpopulations present within the bulk population (Fig. 3.2b). Of these subclusters, cells contained within Cluster 7 constituted the smallest fraction of the bulk population (5.36%) leading us to question if this rare group of cells was functionally distinct from the remaining tumor population. Interestingly, gene set enrichment analysis (GSEA) of the differentially expressed genes from cells identified in Cluster 7 showed a significant enrichment for genes upregulated in basal mammary epithelial cells[32–34] which was surprising given that MCF7 cells are a luminal ER+ model of breast cancer ($P$ < 0.001, Fig. 3.2c, Supplementary Dataset 3.3). Furthermore, there was an increased frequency of cells in Cluster 7 positive for basal candidate markers (e.g. KRT17+, KRT16+) and negative for luminal markers (e.g. KRT18-, KRT19-) suggesting the presence of a basal-like subgroup of cells within this bulk population (Fig. 3.2d).

Differences in gene expression measured from a bulk tumor sample could be due to differences in cell type composition ratios, a version of Simpson's paradox[35]. We therefore hypothesized that using molecular signatures which are variable between patients could aid in the interpretation of differences between single cancer cells. For this purpose we used expression signatures reflective of between-patient heterogeneity in terms of breast cancer subtype[36], and gene modules capturing patient variability derived from a cohort of 342 breast cancer patients[37]. Scoring individual MCF7$^{ER+}$ cells based on their PAM50 gene signature reflected cells with differences in annotated subtype. These differences were also reflected in the single cell expression of gene modules annotated to represent differences in molecular programs present in bulk breast cancer samples (e.g. proliferation, ESR1/ER+, ERBB2/HER2, basal, epithelial-to-mesenchymal transition/EMT programs) (Fig. 3.2e, Supplementary Dataset 3.4). While we found that ESR1

module gene expression was evenly distributed across all MCF7 subpopulations, the basal module gene signature was distinctly elevated in cells specific to Cluster 7 (Fig. 3.2e,f). Taken together, our results indicate the presence of a basal-like subpopulation in an established luminal breast cancer cell line.

We next investigated the degree of functional heterogeneity within another subtype of breast cancer by performing a similar characterization of cells from the HER2-amplified cell line $SKBR3^{HER2}$ (Fig. 3.3a). While expression of *ERBB2* and the ERBB2/HER2 module was consistent across all $SKBR3^{HER2}$ subpopulations (Fig. 3.3b,c), we identified a small group of cells (6.03% of bulk population) comprising Cluster 8 that had a significantly higher frequency of cells positive for EMT markers *CD44* and *FN1* (Fig. 3.3d). Assessing individual cells based on expression of EMT module genes showed a rare population of cells specific to Cluster 8 were highly enriched for this program, consistent with the increased frequency of EMT candidate markers at the single gene level (Fig. 3.3e). Previous studies have shown EMT status can alter drug responses and has been associated with resistance to targeted therapies including HER2 inhibitors[38,39]. In line with these reports, we found that breast cancer cell lines[40,41] with a high expression of EMT module genes were significantly more resistant to the dual HER2/EGFR inhibitor lapatinib (Supplementary Fig. 3.3). As such, we hypothesized that the rare subpopulation of EMT module-high expressing cells observed within the $SKBR3^{HER2}$ cell line may show similar patterns of resistance when treated with lapatinib, despite their HER2 status. To test this hypothesis, we developed a single-live-cell tracking workflow to directly assess the responses of different subpopulations to drug treatment through a competition assay established within a single model system (Fig. 3.3f, Supplementary Fig. 3.4). Using CD44 as a marker to

isolate EMT-high cells, bulk SKBR3$^{HER2}$ cells were separated into CD44$^{hi}$ (top 5%) and CD44$^{lo}$ (remaining) subpopulations using fluorescence-activated cell sorting (FACS). To track the proliferation of CD44$^{hi}$ and CD44$^{lo}$ cells once they are mixed back into a single population, CD44$^{hi}$ cells were labeled with a red cell stain, which was stable over multiple cell divisions throughout a 96 h timeframe (Supplementary Fig. 3.4). In this competition assay following 72 h of lapatinib treatment, we observed higher relative proliferation and decreased apoptosis in CD44$^{hi}$ cells compared to the CD44$^{lo}$ fraction (Fig. 3.3g-i). Hence, using single-cell RNAseq we identified a subpopulation of pre-existing mesenchymal cells within the HER2-amplified SKBR3 cell line that display intrinsic resistance to HER2 inhibition.

### 3.3.3 A subpopulation of cells in an interferon-stimulated gene (ISG)-state is recurrent across TNBCs and associates with worse prognosis

We next sought to characterize the cellular variability in various models of TNBC. Surprisingly, analysis of gene expression profiles from the cells of two TNBC cell lines (HCC38, MDA-MB-231) and a patient-derived xenograft model of TNBC (HCI-002) identified a small subpopulation enriched for genes involved in interferon signaling present in every model tested (HCC38: Cluster 6, 3.86%; MDA-MB-231: Cluster 5, 2.92%; HCI-002: Cluster 1, 37.9%) (Fig. 3.4a). Analysis of the top 50 differentially expressed genes from these subpopulations showed a significant enrichment for genes and gene sets involved in interferon signaling and viral processes, including selective enrichment for interferon responsive genes (*IFIT1*, *IFIT2* and *IFIT3*) and genes involved in antigen presentation (*HLA-A*, *HLA-B*, and *HLA-C*) in HCC38 Cluster 6 cells (Fig. 3.4b,c). Previous work in our lab identified a similar molecular program

representing an immune-like gene set involving many of these genes from breast cancer patient samples[37] (Supplementary Fig. 3.5). Independent studies have also identified many of these genes as a co-expression module in a separate breast cancer cohort[42,43]. Scoring of individual HCC38 and MDA-MB-231 cells showed Cluster 6 and Cluster 5 cells were highly enriched for the ISG module signature, respectively (Fig. 3.4d,e). Looking at the transcript profiles of breast cancer samples from The Cancer Genome Atlas[44] (TCGA, *n* = 342 samples with full data) showed variable expression of ISG module genes across patients, which was similar in structure to the variability observed between individual TNBC cells (Fig. 3.4f,g). Furthermore, there was no enrichment for ISG module genes seen across any subclusters of $SKBR3^{HER2}$ or $MCF7^{ER+}$ cells (Fig. 3.4h). Analysis of individual patient samples from the METABRIC breast cancer cohort[45] (*n* = 1,966 samples) showed a significant enrichment for ISG genes specific to TNBC patients compared to those classified as receptor positive suggesting this phenomenon is specific to the TNBC subtype (Fig. 3.4i). Importantly, survival analysis of TNBC patients from the METABRIC cohort revealed that this recurrent ISG signature associates with a worse prognosis, suggesting patients with a high proportion of cells in this ISG-state may benefit from eradication of this subpopulation (Fig. 3.4j).

### 3.3.4 ISG-state cells display heightened STING-pathway activation and genomic instability

We next investigated the potential mechanisms underlying the presence of a recurrent subpopulation of ISG-cells in TNBCs. Previous work has shown that sustained expression of ISG genes can result from chronic cGAS/STING-pathway activation[42,46] (Fig. 3.5a). One

potential route of STING-pathway activation can be attributed in response to cytosolic DNA fragments, whereby cGAS surveillance of micronuclei links genome instability to innate immunity through TBK1 signaling leading to ISG gene expression[47]. Additionally, the RNA sensing RIG-I/MAVS signaling pathway also converges on TBK1 and can serve as an alternate route to expressing ISG genes[48–50] (Fig. 3.5a). Given that TNBCs are known to have higher levels of genomic instability[44,51–54], we therefore hypothesized that the recurrent subpopulations observed in TNBC may be a result of an ISG-state caused by chronic cGAS/STING activation. Expression of STING (*TMEM173*) and STING effector genes (*IRF3, ISG15, CCL5*) was elevated in HCC38 cells in Cluster 6, which was previously identified as an ISG-module high subpopulation (Fig. 3.5b, 3.4d). Leveraging the fact that hyperactive ISG signaling leads to upregulation of antigen presentation machinery on the cell surface[52], we isolated this subpopulation based on expression of HLA-A/B/C proteins using FACS and found $HLA^{hi}$ cells had increased amounts of total and phosphorylated TBK1 protein, confirming heightened STING-pathway activation in this group of cells (Fig. 3.5c, Supplementary Fig. 3.6). Furthermore, analysis of the dynamics of the HLA subpopulations revealed these cells exist in a transient state (Supplementary Fig. 3.6). We next sought to delineate the cause of activation of this pathway, which could be due to interferon stimulation, viral infection or genomic instability[46]. TNBCs have been previously shown to have increased amounts of genomic instability[44,51–54] which can in theory lead to increased amounts of cytosolic DNA and subsequent activation of STING signaling. We observed a higher number of micronuclei formation in HCC38 $HLA^{hi}$ cells compared to the $HLA^{lo}$ subpopulation, suggesting genomic instability as the potential source of STING-pathway activation in these cells (Fig. 3.5d). Additionally, HCC38 cells identified in the ISG-state showed increased copy number variations again

pointing towards genomic instability as a potential cause of heightened STING-pathway activity and subsequent ISG expression (Supplementary Fig. 3.7).

### 3.3.5 TNBC cells in the ISG-state are chemoresistant

Since patients with high expression of an ISG signature were associated with worse prognosis (Fig. 3.4j), we next investigated the functional consequence a subpopulation of cells in the ISG-state may have on treatment responses. In order to predict the response of these cells to therapy, we performed pharmacogenomics modeling based on bulk breast cancer cell line drug sensitivities. First, a panel of 82 molecularly characterized breast cancer cell lines[40,41] were scored based on their ISG-module gene expression. Drug response data covering 90 compounds for these cell lines was then correlated with ISG module scores to identify compounds associated with sensitivity and resistance in the bulk breast cancer cell lines (Fig. 3.6a). We found that cell lines with high ISG module scores were significantly associated with resistance to the chemotherapies gemcitabine (r = -0.4244, $P$ = 0.0079) and doxorubicin (r = -0.3054, $P$ = 0.05) (Fig. 3.6b). We therefore predicted that subpopulations of ISG cells would show similar patterns of chemoresistance, and tested this using our single-live-cell tracking workflow to assess the sensitivities of HCC38 $HLA^{hi}$ (top 5%) and $HLA^{lo}$ (remaining) subpopulations in response to chemotherapy (Supplementary Fig. 3.8). As predicted, HCC38 $HLA^{hi}$ cells were resistant to gemcitabine compared to $HLA^{lo}$ cells as evidenced by increased proliferation, decreased apoptosis, and an increase in fraction of $HLA^{hi}$ cells in the remaining population following a 72 h drug treatment (Fig. 3.6c-e). We observed a similar decrease in sensitivity in HCC38 $HLA^{hi}$ compared to the $HLA^{lo}$ subpopulation following treatment with doxorubicin confirming cells in

a ISG-state are resistant to these two chemotherapies (Fig. 3.6f-h). Furthermore, we confirmed a subpopulation of ISG cells in a second model of TNBC (MDA-MB-231) also had increased survival in response to gemcitabine and doxorubicin following long-term drug treatment (Supplementary Fig. 3.9). Consistent with our model, several studies have identified interferon-related gene signatures upregulated upon drug resistance and in models of residual disease in breast cancer[14,48,55,56] which may occur due to selection for this population by chemotherapy as we observed *in vitro*.

### 3.3.6 Eliminating chemoresistant subpopulations by targeting tumor cells in an ISG-state with CDK4 or mTOR inhibitors

We next asked if we could target this chemoresistant subpopulation of chronically inflamed TNBC cells using the predictions made from our pharmacogenomic modeling of bulk breast cancer cell line data. Our analysis identified cells in an ISG-state as hypersensitive to CDK4 inhibitor fascaplysin and the mTOR inhibitor rapamycin suggesting a potential mode to target this chemoresistant subpopulation pharmacologically (Fig. 3.6a). We confirmed this effect *in vitro* by isolating $HLA^{hi}$ and $HLA^{lo}$ subpopulations of HCC38 cells and found decreased survival of $HLA^{hi}$ cells in response to long term treatment with palbociclib (CDK4i), ribociclib (CDK4/6i) and everolimus (mTORi) (Fig. 3.7a-c). Notably, whereas single-agent gemcitabine treatment could not robustly induce cell death in $HLA^{hi}$ cells, the addition of palbociclib in combination showed an increase in cell death on both subpopulations suggesting a potential strategy for targeting the full tumor cell population (Fig. 3.7d-f). Supporting this, analysis of subpopulations within a single mixed model of HCC38 cells showed single-agent gemcitabine or

palbociclib selectively induced apoptosis in $HLA^{lo}$ or $HLA^{hi}$ cells respectively, while combination treatment resulted in increased cell death in both subpopulations (Fig. 3.7g). Together this data suggests targeting TNBC cells in an ISG-state with palbociclib can increase the efficacy of chemotherapy treatments by eradicating a resistant subpopulation leading to increased elimination of the full tumor landscape.

## 3.4 Discussion

Through pharmacogenomic modeling based on bulk cancer cell line drug sensitivities, we developed a new approach using a Modular Drug Response Prediction Pipeline ($MOD_{Rx}$) in order to systematically discover drug combinations capable of counteracting tumor heterogeneity by predicting drug responses of individual breast cancer cells. Analysis of single-cell profiles from breast cancer cell lines and patient-derived xenograft models in conjunction with data from patient tumor samples revealed a high magnitude of heterogeneity both across different lineages and within each model system analyzed. Inference of CNV levels of individual cells showed multiple genetic subclones are present within transcriptionally distinct subpopulations, corroborating external reports showing some but not all variability between cancer cells can be attributed to subclonality[9,24].

We reasoned that the molecular variability between individual breast cancer cells and subsequent impact on drug sensitivities could be modeled by taking an unbiased modular analysis approach

that incorporates differences observed between breast cancer patients and may better capture and explain the heterogeneity seen between cells. In doing so, we identified a rare subpopulation of basal-like cells within the MCF7 ER+ cell line, in line with previous studies showing breast cancer cell lines are populated by cells coexisting in numerous differentiation-states with distinct luminal, basal, and mesenchymal gene expression patterns leading to heterogeneity in these model systems[16,57]. This finding has severe clinical implications in that targeted therapies have been shown to enrich for distinct differentiation states in breast cancers. For example, treatment with the MEK inhibitor trametinib has been shown to enrich for breast cancer cells in a basoluminal (K19/K5/K14-high, K8/VIM-low) state suggesting target therapies have different impacts on cell lineages and drug combinations are required to reduce therapeutic escape[16]. Building on this, we developed a computational approach to identify signatures of drug resistance and sensitivity of individual cells based on bulk cell line pharmacologic profiling datasets and patient-derived gene modules and applied this approach in the context of the HER2-amplified cell line SKBR3. We identified a subgroup of pre-existing mesenchymal cells within the SKBR3 population that displayed intrinsic resistance to lapatinib which we experimentally validated using a single-live-cell tracking workflow.

Our data identified a recurrent subpopulation of cells in an ISG-state that was specific to TNBC cells and displayed heightened STING-pathway activation and genomic instability. By applying our $MOD_{Rx}$ approach we determined these cells are associated with chemoresistance and poor outcome in TNBC patients. This finding is aligned with previous studies defining interferon-related gene signatures upregulated upon drug resistance and potentially explains the cellular source of heightened ISG expression found in residual TNBC tumors after

chemotherapy[14,48,55,56]. Furthermore, our analysis predicted ISG-state cells as hypersensitive to CDK4/6 and mTOR inhibition, which we validated as an effective strategy for targeting this chemoresistant subpopulation in TNBC. Guided by our *in silico* modeling and single-cell predictions, we show the addition of palbociclib to gemcitabine or doxorubicin treatment is an effective strategy for increasing therapeutic efficacy on TNBC tumor landscapes in their entirety.

By taking a modular approach to assessing breast cancer heterogeneity we were able to characterize molecular and functional differences between transcriptionally distinct subsets of cells and ultimately link tumor transcriptional heterogeneity with drug actionability. There were several reasons for approaching this study using $MOD_{Rx}$: (1) ideal use-case for studying single-cells in that aggregating genes into modules facilitates processing the high-dimensional data acquired from sequencing at the single-cell level; (2) using modules overcomes technical challenges associated with analyzing single-cell expression profiles due to sparsity in mRNA measurements and provides a method for detecting zero-inflated genes in single-cell transcriptomics data; (3) modules simplifies gene sets making analysis more computationally- and time-efficient; (4) gene modules are effective tools for connecting functional characteristics to individual cells in that they are representative of co-expression patterns in patients, provide pathway-enrichment insight, and can be correlated with drug response data from large-scale databases. This affords a useful tool for rapidly defining drug sensitivities and resistance patterns within tumors by bypassing the need for extensive functional screening of each independent subpopulation for candidate identification, and ultimately leading to quicker clinical translation.

Using *in silico* and experimental approaches, our study provides an effective new framework to discover drug combinations capable of counteracting tumor heterogeneity by predicting drug responses of individual breast cancer cells and systematically links tumor transcriptional heterogeneity with drug actionability to optimize therapy combinations. Ongoing efforts are currently underway to translate this approach at scale by developing a machine-learning model for generating a drug-by-cell matrix to provide a systematic method for predicting single-cell responses and inform candidate combination strategies.

## 3.5 Methods

### 3.5.1 Breast cancer cell lines and reagents

SKBR3$^{HER2}$ cells were obtained from the UCSF Cell Culture Facility. HCC38$^{TNBC}$, MCF7$^{ER+}$, and MDA-MB-231$^{TNBC}$ cell lines were obtained from the American Type Culture Collection (ATCC). Cell lines were grown according to published protocols[58] except for SKBR3 cells, which were cultured using RPMI media supplemented with 10% fetal bovine serum (FBS) and 1% pen/strep. All cell lines tested negative for mycoplasma contamination. Drugs used for cell culture experiments in this study were purchased from Selleck Chemicals (Gemcitabine, Doxorubicin, Ribociclib) and LC Laboratories (Palbociclib, Everolimus).

### 3.5.2 Patient-derived xenograft model and tissue dissociation

The HCI-002$^{PDX\text{-}TNBC}$ patient-derived xenograft (PDX) tumor sample was flash frozen in liquid nitrogen following a previous study where tumors were grown as previously described[59,60]. Viable tumor sample was then thawed and dissociated for fluorescence-activated cell sorting (FACS) using established protocols[61]. Briefly, tissue was mechanically chopped with scalpels and placed in culture medium DMEM/F12 with 5% FBS, 5 μg $ml^{-1}$ insulin (UCSF Cell Culture Facility), 50 ng $ml^{-1}$ gentamycin (UCSF Cell Culture Facility) and supplemented with 2 mg $ml^{-1}$ collagenase-1 (Sigma). Sample was then digested for 45 min at 37 ˚C. The resulting suspensions were resuspended in 2 U $\mu l^{-1}$ DNase (D4263-5VL, Sigma Aldrich) for 3 min at room temperature, washed and dissociated with 2 ml of 0.05% trypsin/EDTA (25-052-CI, Corning) for 10 min at 37 ˚C. Cell suspensions were then filtered through a 70 μm filter, and frozen in DMEM/F12 with 50% serum, 10% DMSO, and stored in liquid nitrogen prior to study.

### 3.5.3 Single-cell RNA sequencing (scRNAseq)

Breast cancer cell lines (MCF7$^{ER+}$, SKBR3$^{HER2}$, MDA-MB-231$^{TNBC}$) were thawed and carried according to published culture methods as described. Viably frozen cell suspensions of PDX tumor sample (HCI-002$^{PDX\text{-}TNBC}$) were thawed and washed using DMEM prior to FACS sorting. Cells were stained for FACS using fluorescently labeled antibodies for human antigen CD298 (PE; 341704, BioLegend) and mouse antigens CD45 (APC; 559864, BD Pharmingen), CD31 (APC; 551262, BD Pharmingen) and Ter119 (APC; 557909, BD Pharmingen). Flow sorting was done using a BD FACSAria II cell sorter (Becton Dickinson) where contaminating

hematopoietic and endothelial cells were excluded by gating out $Lin^+$ (CD45, Ter119, CD31) cells, and dead cells were eliminated by excluding Sytox positive (SYTOX Blue Dead Cell Stain, S34857, Life Technologies) cells. Samples showing at least 80% viability as measured using SytoxBlue in FACS were proceeded with for single-cell sequencing.

For droplet-enabled scRNAseq, flow cytometry sorted cells were washed in PBS with 0.04% BSA and resuspended at a concentration of ~1000 cells/μl. Single-cell RNA sequencing was performed under the supervision of the IHG Genomics Core (UCSF) using the Chromium Single Cell 3′ Reagents Kit (CG00026 Rev B., 10x Genomics), and libraries were prepared following the manufacturer's protocol. Quantification of cDNA libraries was performed using Qubit dsDNA HS Assay Kit (Q32851, Life Technologies) and high-sensitivity DNA chips (5067-4626, Agilent). Libraries were then sequenced using the Illumina HiSeq2500 platform to achieve an average calculated depth of 50,000 reads per cell.

### 3.5.4 scRNAseq data processing and analysis

The Cell Ranger Single-Cell Software Suite v1.1.0 was used to perform sample demultiplexing, barcode processing and single-cell 3′ gene counting (http://software.10xgenomics.com/single-cell/overview/welcome). For cluster identification in the scRNAseq datasets, we utilized the Seurat pipeline[27,28,30]. The data matrices were imported into R and processed initially with the Seurat R package version 2.0 and later with Seurat R package version 3.0, where the FPKM values were transformed into log-space after the aforementioned trimming steps (each gene was expressed in at least three cells, each cell has at least 200 genes). Principal component analysis

(PCA) was performed using highly variable genes for each separate sample in the trimmed dataset. Using significant principal components (determined by Jackstraw method) as input, we then performed density clustering to identify groupings in the data and *t*-distributed statistical neighbor embedding (tSNE) to visualize. Marker genes for each respective cluster were identified and utilized in subsequent analysis using further Seurat functionality.

### 3.5.5 Module scoring

Molecular programs describing breast cancer biology and representing breast cancer patient variability were previously defined using MAGNETIC[37]. To compare the identified 219 gene signatures in subpopulations, gene module scores were calculated for each individual cell in order to evaluate the degree to which single-cells express a given module. Of note, cells with higher complexity would be expected to have higher cell scores for any set of genes analyzed. To account for this confounding effect, for each gene module analyzed, a control gene-set was created to be used in the calculation of a normalization factor as described previously[62]. Briefly, control gene sets were selected to ensure a similar distribution of expression levels to that of the input gene module. First, all genes analyzed are ordered by mean expression across all cells and partitioned into 25 bins of similar expression. Next, for each gene in the given gene module or signature, 5 control genes were randomly selected from its corresponding expression bin and added to a control set. The resulting control set contained an equivalent expression distribution as the target gene module and its average represents an equivalent sampling of 100 gene sets of equal size to the target gene module. The expression of genes in the target gene module and the control gene set was averaged across each cell to generate a target score ($S_{Target}$) and control score ($S_{Ctrl}$).

The cell's score for the target gene module ($S_{Mod}$) is then the difference between the target score and control score: $S_{Mod} = S_{Target} - S_{Ctrl}$.

### 3.5.6 Gene Set Enrichment Analysis (GSEA)

Gene set enrichment analysis (GSEA) of hallmark cancer gene signatures in the Molecular Signatures Database (MSigDB v6.0) was performed using GSEA v3.0 software (http://www.broadinstitute.org/gsea/)[33,34] under the following parameters: permutation, phenotype; metric, Signal2Noise; scoring scheme, weighted; and number of permutations, 1,000. Gene sets were considered significantly enriched following a nominal $P < 0.05$ and FDR < 0.25 cutoff. Cells belonging to subpopulations were averaged to serve as a representation of each subgroup, and trimmed to their respective marker genes as determined by Seurat following log transformation.

### 3.5.7 Fluorescence-activated cell sorting (FACS) of subpopulations

Excluding FACS analysis prior to sequencing as described above, sorting of SKBR3, HCC38, and MDA-MB-231 cells was performed on a Sony SH800S Cell Sorter (Sony) using the following fluorochrome-conjugated antibodies at the manufacturer's recommended concentration unless otherwise indicated: anti-human CD44 conjugated to APC-H7 (560532, BD Pharmingen) at 1:200, and anti-human HLA-A/B/C conjugated to Alexa Fluor® 488 (560169, BD Pharmingen). Gating of positive and negative cells was defined by the unstained control, and cells were sorted into representative high and low expressing populations as indicated. For ISG module kinetics experiments, HCC38 cells of each subpopulation

(ISG/HLA$^{hi}$: top 5%; ISG/HLA$^{lo}$: bottom 10%, control sort: bulk HCC38) were sorted and reanalyzed by FACS immediately post-sorting and at 14 d afterwards. Final analysis was performed using FlowJo Software v10.6.1 (Tree Star).

### 3.5.8 Single-live-cell tracking

We developed a single-live cell tracking workflow to directly assess the functional implications of various subpopulations of cancer cells within a heterogeneous mixed-model system. First, subpopulations of interest were isolated using FACS as previously described. Following sorting, we then labeled the cytoplasm of living cells from subpopulations of interest using the IncuCyte® CytoLight Rapid Red Reagent (4706, Essen BioScience) per the manufacturer's protocol, while additional subpopulations were left unlabeled. This stably fluorescent dye freely passes through cell membranes and transforms into a cell-impermeant fluorescent probe that is transferred by dilution to daughter cells without affecting adjacent cells within a population, allowing for tracking subpopulations of cells over time without perturbing cell function or biology. Reagent concentration was independently optimized in all breast cancer cell lines used in the study to determine fluor stability duration and detection limits for monitoring cells in mixed cultures. IncuCyte® Caspase-3/7 Green Reagent (4440, Essen BioScience) or YO-PRO-1 (Thermo Fisher Scientific) was multiplexed with the tracking reagent to simultaneously determine apoptotic or cytotoxic effects within wells. Both labeled and unlabeled subpopulations of cells were then mixed back into a homogenous system and seeded in 384-well assay microplates and allowed to adhere over 8 h. Cells were then treated with drug and imaged over 72 h at 2 h timepoints using the IncuCyte® S3 Live-Cell Analysis System with a Green/Red Optical

Module. Capturing images from phase, green, and red channels allowed for monitoring unlabeled cells, apoptotic cells, and labeled subpopulations respectively.

### 3.5.9 Drug sensitivity studies

Cells were seeded in 384-well assay microplates at a density of 1,000 cells/well in a total volume of 40 μL/well and incubated at 37°C, 5% $CO_2$ overnight. Following 72 h drug exposure, proliferation and cell death was measured by staining with Hoechst 33342 (Thermo Fisher Scientific) nuclear dye and YO-PRO-1 (Thermo Fisher Scientific) early apoptosis dye, respectively. Stained plates were then analyzed using a CellInsight High Content microscope (Thermo Fisher Scientific). Raw phenotype measurements from each treated well were normalized to the median of vehicle-treated control wells and examined for response.

### 3.5.10 Colony formation assay

Colony formation assays were performed using crystal violet staining and quantification. Briefly, cells were seeded in 12-well microplates at a density of 1,000 cells/well and incubated at 37°C, 5% $CO_2$ overnight. Cells were then exposed to drug or DMSO vehicle for 10-14 d, with medium change and drug refresh every 3 d. Cells were fixed with 4% formaldehyde and stained with 0.5% crystal violet. Images of stained cells were taken using an EPSON Perfection V600 scanner prior to quantification. Growth was quantified by dissolving crystal violet in 0.1% SDS and absorbance was measured at 590 nm using a spectrophotometer and normalized to vehicle-treated controls.

### 3.5.11 Immunoblot

Subpopulations of cells were collected via FACS and proteins were extracted using RIPA buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% sodium deoxycholate, 0.1% SDS, 1 mM EDTA pH 8.0, 1% NP-40) supplemented with protease and phosphatase inhibitor cocktails (Roche). Samples were sonicated and then centrifuged at 14,000 r.p.m. for 15 min at 4°C. Protein concentrations were determined by Bradford assay (Bio-Rad). Equal amounts of protein samples were resolved using 4–12% SDS-PAGE gels (Life Technologies) and transferred to polyvinylidene difluoride membranes (Millipore). Membranes were probed overnight on a 4°C shaker with primary antibodies (1:1,000 dilution unless indicated) recognizing the following proteins: p-TBK1 (Ser172) (5483, Cell Signaling), TBK1 (3504, Cell Signaling), cGAS (15102, Cell Signaling), p-S6 (Ser240/244) (5364, Cell Signaling, 1:20,000), p-4E-BP1 (Thr37/46) (2855, Cell Signaling), p-STAT1 (Tyr701) (9167, Cell Signaling), MYC (ab32072, Abcam), and β-actin (3700, Cell Signaling). Proteins were detected via incubation with horseradish peroxidase-conjugated secondary antibodies, Clarity Western ECL Blotting Substrates (Bio-Rad) or SuperSignal West Femto Maximum Sensitivity Chemiluminescent Substrate (Thermo Fisher Scientific).

### 3.5.12 Immunofluorescence

After sorting subpopulations via FACS, cells were grown on glass coverslips and fixed with 4% paraformaldehyde (PFA) in PBS for 10 min at room temperature. Cells were then permeabilized in 1X PBS/0.3% Triton X-100 for 10 min at room temperature before blocking for 40 min with 3% BSA in PBS. Coverslips were then incubated with primary antibody overnight at 4°C,

followed by incubation with a secondary antibody for 1h at room temperature. Both primary and secondary antibodies were diluted in blocking buffer given the following dilutions: cGAS (D1D3G) (Cell Signaling, 1:200), anti-rabbit-Alexa 488 (A11008) (Invitrogen, 1:1,000). Coverslips were then mounted using Vectashield Antifade Mounting Medium with DAPI (H-1200, Vector Laboratories) and imaged using a Zeiss LSM 780 confocal microscope with 25× and 63× objectives. The images were further processed in ImageJ and scoring was performed under blinded conditions.

### 3.5.13 Quantification of micronuclei frequency

After fixation and DAPI staining, cells containing micronuclei were visualized by microscopy using a confocal microscope (Zeiss LSM 780) equipped with a 25× and 63× water immersion objective and a CCD camera. Cells were analyzed under blinded conditions and micronuclei positive fractions calculated as a percentage of total cells per field. For quantification, multiple (3-5) random fields were captured and 700-1000 cells were counted in each independent experiment. Micronuclei were defined as discrete DNA aggregates separate from the primary nucleus in cells where interphase primary nuclear morphology was normal. Cells displaying mitotic morphology and/or with an apoptotic appearance were excluded.

### 3.5.14 CNV estimation based on single-cell RNAseq data

Large-scale copy number variations (CNVs) were inferred from single-cell expression data using the InferCNV algorithm[31]. Initial CNVs were estimated by sorting the analyzed genes by their

chromosomal location and applying a moving average to the relative expression values, with a sliding window of 100 genes within each chromosome, as previously described[63].

### 3.5.15 Statistical analysis

Data are expressed as means ± s.d., unless otherwise indicated. Statistical analyses were performed using GraphPad Prism 6 (v6.0g) and R (v3.4.4). Two-tailed Student *t*-tests (with unequal variance) were used in all comparisons unless otherwise noted. $P < 0.05$ was considered statistically significant throughout the study.

### 3.5.16 Data availability statement

All data generated or analyzed during this study are included in this published article and its supplementary information files.

## 3.6 Acknowledgements

The authors would like to thank members of the Bandyopadhyay laboratory for helpful discussions and technical assistance. We also acknowledge the IHG Genomics core and the Center for Advanced Technology (CAT) core and staff at the University of California San Francisco for sequencing support. We also thank A. Zaman, H. Goodarzi, and A. Paulson for helpful discussions and reagents. This work was supported in part by NCI U01CA168370 (S.B.), NIGMS R01GM107671 (S.B.), Martha and Bruce Atwater (S.B.), and Prospect Creek Foundation (S.B., A.G.).

## 3.7 Author contributions

H.J.D. and S.B. contributed toward study conceptualization. H.J.D., S.K., and A.T. performed data analyses supporting the study. H.J.D. designed and performed the majority of experiments. O.M. assisted with samples for initial single-cell sequencing. H.J.D. and S.B. composed the original draft and all authors contributed to manuscript finalization. S.B. and A.G. supervised the study.

# 3.8 Figures

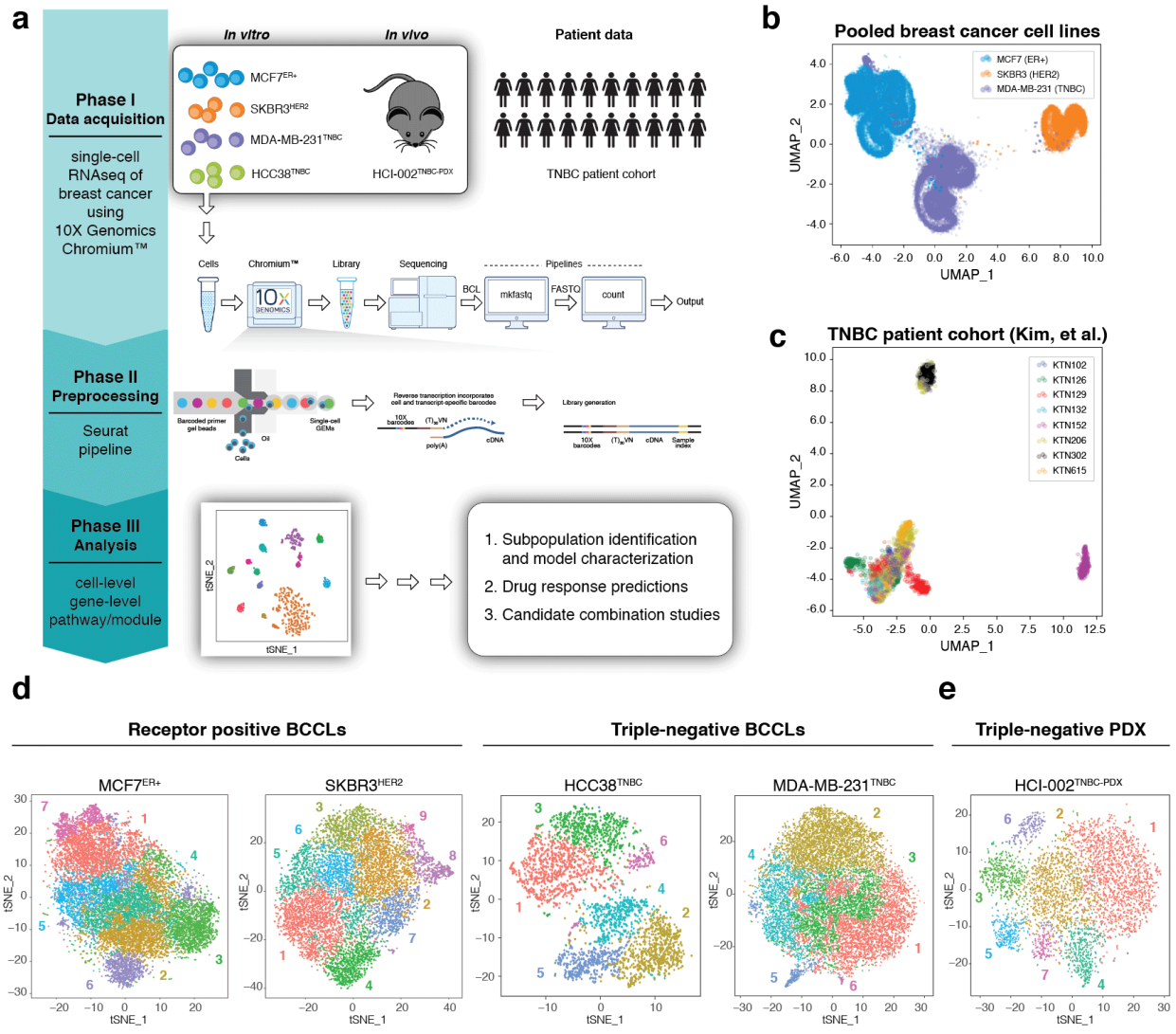



**Figure 3.1 Single-cell RNAseq across multiple models of breast cancer.** (**a**) Schematic of single-cell profiling approach. (**b**) Combined UMAP projection of cells from various breast cancer cell lines representing different ER+, HER2-amplified, and TNBC subtypes. Individual cells are colored based on cell line identity. (**c**) UMAP projection of cells from a cohort of 8 TNBC patients[9] Individual cells are colored based on patient identity. (**d**) tSNE map of cells from breast cancer cell lines. (**e**) tSNE map of cells from a patient-derived xenograft model of TNBC (HCI-002$^{\text{PDX-TNBC}}$). For (**d,e**) individual cells are colored by clusters identified using independent graph-based clustering analysis for each model tested. BCCL, breast cancer cell line; TNBC, triple-negative breast cancer; PDX, patient-derived xenograft.

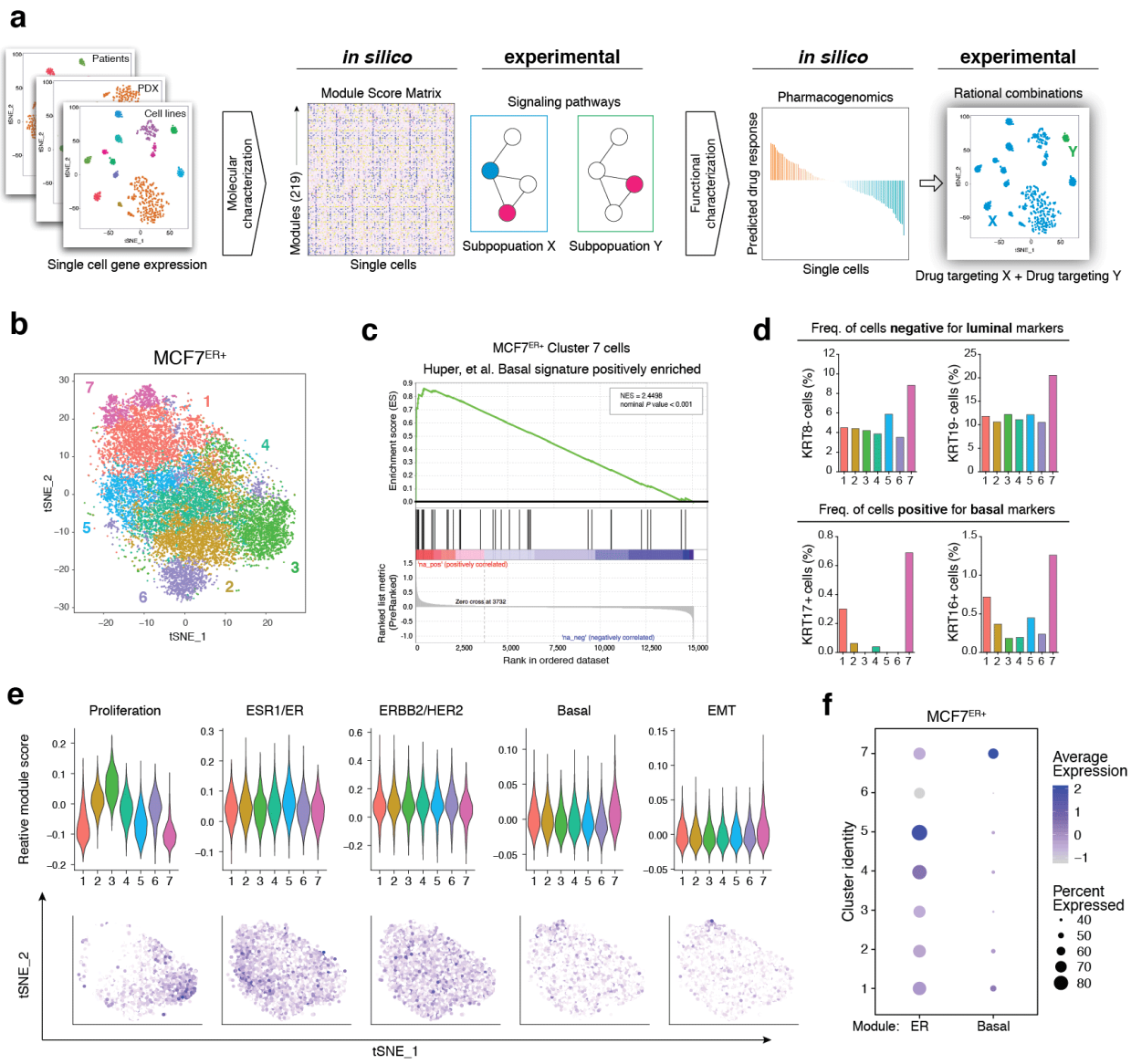



**Figure 3.2 Modular analysis of individual breast cancer cells informs functional heterogeneity and mirrors inter-patient variability.** (**a**) Schematic for designing rational combination strategies using a Modular Drug Response Prediction Pipeline ($MOD_{Rx}$). (**b**) tSNE plot showing subclusters of cells identified within a $MCF7^{ER+}$ bulk population. (**c**) Gene set enrichment analysis (GSEA) of transcript profiles from $MCF7^{ER+}$ cells identified in Cluster 7 compared to non-Cluster 7 cells. (**d**) Frequency of $MCF7^{ER+}$ cells negative for luminal markers (KRT18, KRT19; top) and positive for basal markers (KRT17, KRT16; bottom) as a percentage of all cells within each subcluster. (**e**) Relative module score for known molecular breast cancer programs (proliferation, ESR1/ER, ERBB2/HER2, basal, EMT) of individual $MCF7^{ER+}$ cells within subclusters. Gene module expression for individual cells mapped onto tSNE plots are shown below. (**f**) Average expression and percentage of cells expressing ESR1/ER and basal module genes for each subcluster of $MCF7^{ER+}$ cells.

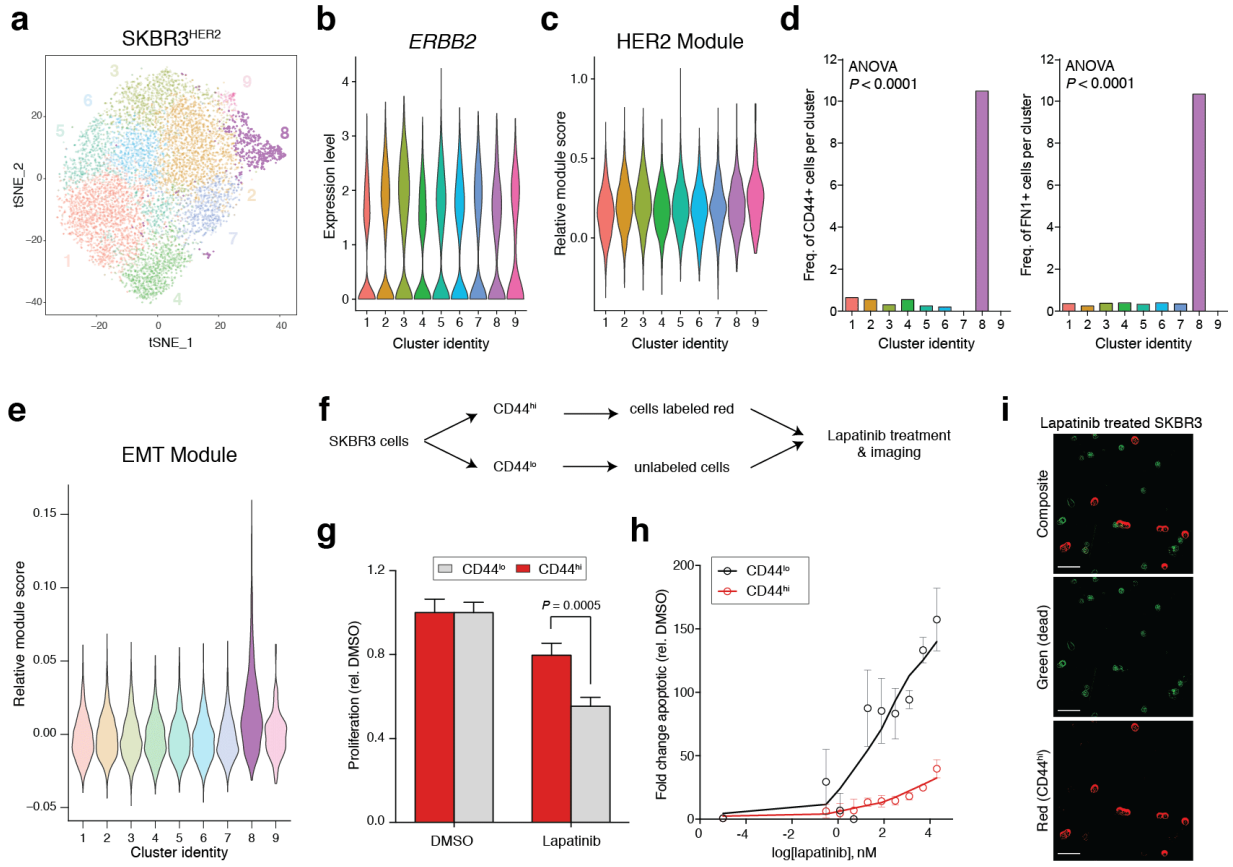



**Figure 3.3 A subpopulation of pre-existing mesenchymal cells within the SKBR3$^{HER2}$ cell line display intrinsic resistance to lapatinib.** (**a**) tSNE plot showing subclusters of cells identified within a SKBR3$^{HER2}$ bulk population, highlighting Cluster 8. (**b**) *ERBB2* expression levels for individual SKBR3$^{HER2}$ cells. (**c**) Relative ERBB2/HER2 module score for individual SKBR3$^{HER2}$ cells. (**d**) Frequency of SKBR3$^{HER2}$ cells positive for markers of EMT (CD44, left panel; FN1, right panel). *P* values calculated using one-way ANOVA. (e) Relative EMT module score for individual SKBR3$^{HER2}$ cells. (**f**) Single-live-cell tracking workflow scheme. (**g**) Relative proliferation of CD44$^{lo}$ (unlabeled) and CD44$^{hi}$ (red) SKBR3 cells following 72 h treatment with 5 μM lapatinib. Data represents $n$ = 4 biologically independent samples. Error bars are mean ± s.d., and *P* values calculated using a two-sided *t*-test. (**h**) Fold change apoptotic cells following 72 h lapatinib treatment for CD44$^{lo}$ and CD44$^{hi}$ SKBR3 cells. Data represents $n$ = 4 biologically independent samples. Error bars are mean ± s.d. (**i**) Representative images of SKBR3 subpopulations (CD44$^{lo}$, unlabeled; CD44$^{hi}$, red) after 72 h treatment with 80 nM lapatinib showing selective cell death in CD44$^{lo}$ cells. Dead cells are labeled green. Images are shown using a 10× objective. Scale bars, 50 μm.

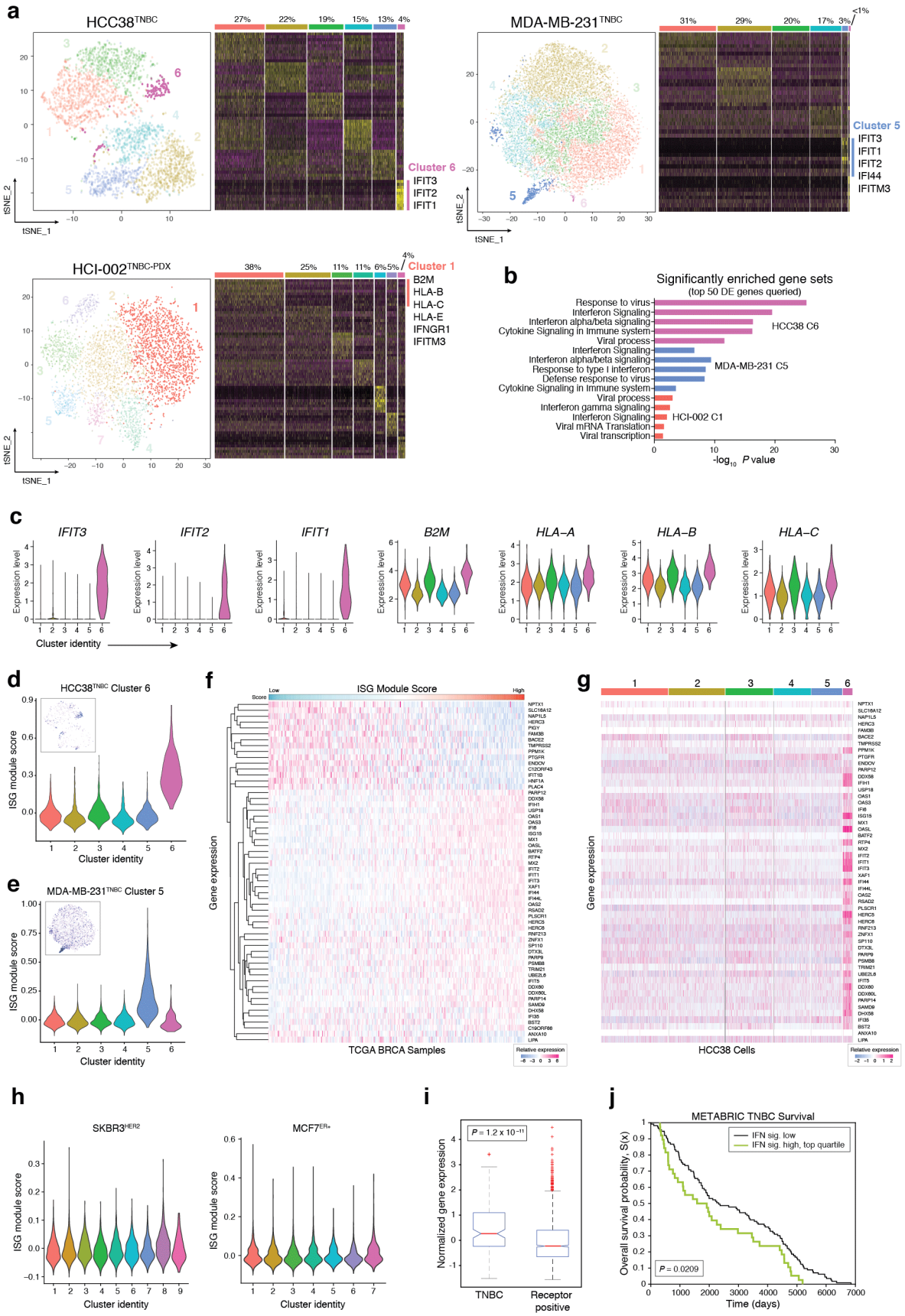

a
HCC38TNBC
27% 22% 19% 15% 13% 4%
Cluster 6
IFIT3
IFIT2
IFIT1
MDA-MB-231TNBC
31% 29% 20% 17% 3% <1%
Cluster 5
IFIT3
IFIT1
IFIT2
IFI44
IFITM3
HCI-002TNBC-PDX
38% 25% 11% 11% 6% 5% 4%
Cluster 1
B2M
HLA-B
HLA-C
HLA-E
IFNGR1
IFITM3
tSNE_1
tSNE_2
b
Significantly enriched gene sets
(top 50 DE genes queried)
Response to virus
Interferon Signaling
Interferon alpha/beta signaling
Cytokine Signaling in Immune system
Viral process
HCC38 C6
Interferon Signaling
Interferon alpha/beta signaling
Response to type I interferon
Defense response to virus
Cytokine Signaling in Immune system
MDA-MB-231 C5
Viral process
Interferon gamma signaling
Interferon Signaling
Viral mRNA Translation
Viral transcription
HCI-002 C1
-log10 P value
c
IFIT3
IFIT2
IFIT1
B2M
HLA–A
HLA–B
HLA–C
Expression level
Cluster identity
d
HCC38TNBC Cluster 6
ISG module score
Cluster identity
e
MDA-MB-231TNBC Cluster 5
ISG module score
Cluster identity
f
ISG Module Score
Low
High
Score
Gene expression
TCGA BRCA Samples
Relative expression
g
1 2 3 4 5 6
Gene expression
HCC38 Cells
Relative expression
h
SKBR3HER2
MCF7ER+
ISG module score
Cluster identity
i
P = 1.2 x 10−11
Normalized gene expression
TNBC
Receptor positive
j
METABRIC TNBC Survival
IFN sig. low
IFN sig. high, top quartile
P = 0.0209
Overall survival probability, S(x)
Time (days)

**Figure 3.4 A subpopulation of cells in an interferon-stimulated gene (ISG)-state is recurrent in TNBCs.** (**a**) tSNE plots for individual cells from two TNBC cell lines (HCC38, MDA-MB-231) and a patient-derived xenograft model of TNBC (HCI-002) highlighting clusters enriched for ISG genes. Heatmaps displaying scaled expression patterns of top marker genes within each cluster shown to the right with ISG genes highlighted (yellow indicates high, and purple indicates low). Percentage of cells contained in each identified cluster is listed at the top. (**b**) Significantly enriched gene sets identified from a Gene Set Enrichment Analysis (GSEA) of the top 50 differentially expressed genes from HCC38 Cluster 6, MDA-MB-231 Cluster 5, and HCI-002 Cluster 1 cells. (**c**) Expression levels of various interferon genes for individual HCC38 cells. Relative ISG-module scores for individual (**d**) HCC38 and (**e**) MDA-MB-231 cells. Insert shows ISG-module expression for individual cells mapped onto tSNE plots. (**f**) Relative expression levels for ISG-module genes across breast cancer patients from the TCGA. Patients sorted based on ISG-module score. (**g**) Relative expression levels for ISG-module genes across individual HCC38 cells. Cells are sorted based on cluster identity listed above. (**h**) Relative ISG-module scores for individual $SKBR3^{HER2}$ (left) and $MCF7^{ER+}$ (right) cells. (**i**) Normalized gene expression levels for ISG genes comparing TNBC and receptor positive breast cancer patients from METABRIC data. Box-and-whisker plots show median (centerline), upper/lower quartiles (box limits), and whiskers spanning the interquartile range from 25-75 percentiles. *P* value calculated using a two-sided *t*-test. (**j**) Kaplan-Meier survival curves of TNBC METABRIC samples stratified based on IFN gene signature status. *P* value calculated using a log-rank test.

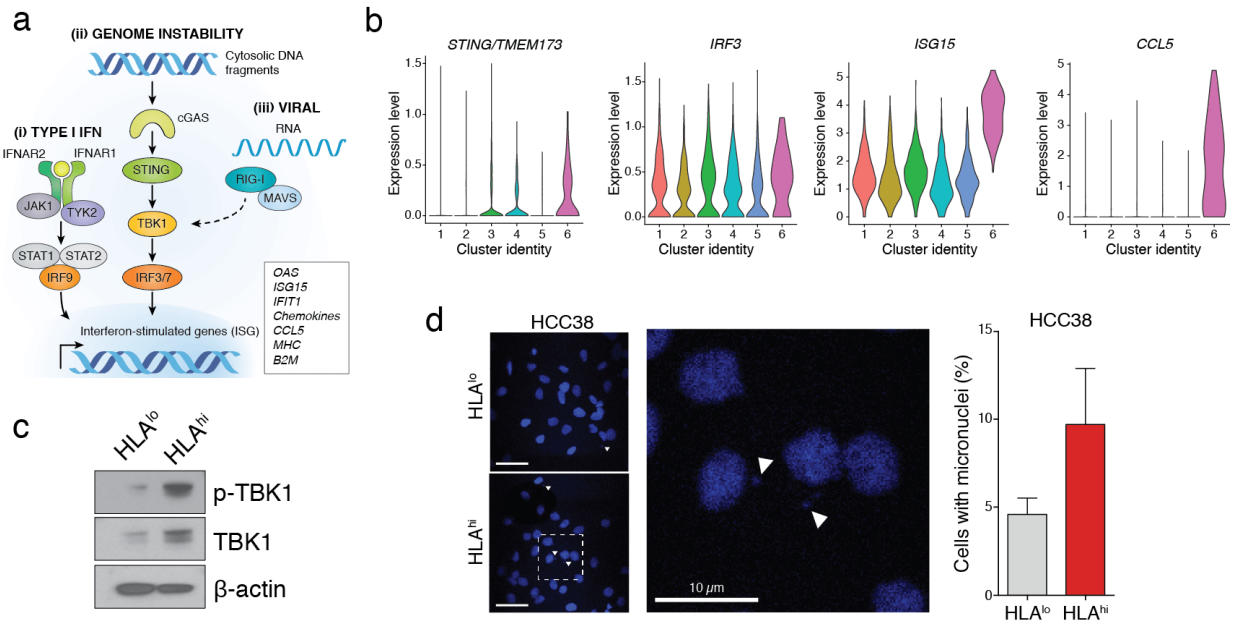



**Figure 3.5 ISG-state cells display heightened STING-pathway activation and genomic instability.** (**a**) cGAS/STING signaling pathway model. (**b**) Expression levels of various STING effector genes for individual HCC38$^{TNBC}$ cells based on cluster identity. (**c**) HCC38 cells were sorted into HLA$^{hi}$ (top 5%) and HLA$^{lo}$ (remaining) populations by FACS, and protein lysates were subjected to immunoblot using the indicated antibodies. Representative image from $n$ = 3 independent experiments. (**d**) Representative images of HCC38 HLA$^{lo}$ (top) and HLA$^{hi}$ (bottom) subpopulations with DAPI (blue) staining DNA. Arrows indicate micronuclei positive cells. Higher magnification of HCC38 HLA$^{hi}$ cells positive for micronuclei shown to the right. Percentage of cells positive for micronuclei is quantified at the right. Data is an average of five high-powered (25×) fields (≥150 cells/field) analyzed per subpopulation. Error bars are mean ± s.d., and $P$ value calculated using a two-sided $t$-test. Images are shown using a 25× objective. Scale bars, 20 μm unless indicated otherwise.

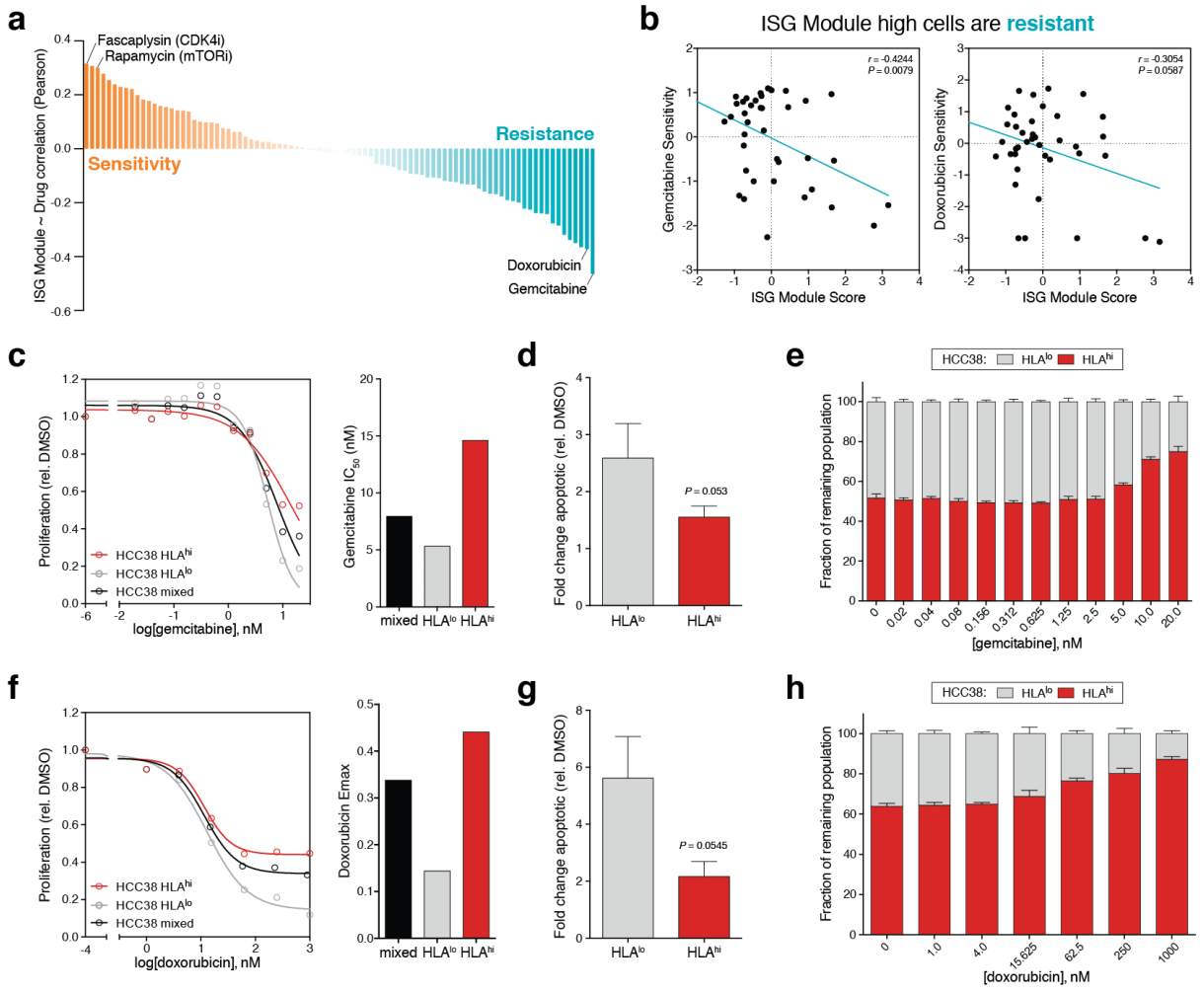



**Figure 3.6 TNBC cells in the ISG-state are chemoresistant.** (**a**) ISG-module scores from a panel of 82 breast cancer cell lines were correlated with drug response values across 90 compounds. Pearson correlation values shown. (**b**) Scatter plot of breast cancer cell line ISG-module scores associated with gemcitabine (left panel; r = -0.4244, $P$ = 0.0079) and doxorubicin (right panel; r = -0.3054, $P$ = 0.05) sensitivities. $P$ values based on Pearson correlations. (**c**) Proliferation of HCC38 HLA$^{lo}$, HLA$^{hi}$, and mixed-bulk populations in response to 72 h gemcitabine treatment. $IC_{50}$ quantification of dose-response curves shown at the right. (**d**) Fold change apoptotic HCC38 cells for HLA subpopulations after 72 h treatment with 5 nM gemcitabine. (**e**) Fraction of remaining cell population consisting of HLA$^{lo}$ and HLA$^{hi}$ HCC38 cells after 72 h gemcitabine treatment. (**f**) Proliferation of HCC38 HLA$^{lo}$, HLA$^{hi}$, and mixed-bulk populations in response to 72 h doxorubicin treatment. Emax quantification of dose-response curves shown at the right. (**g**) Fold change apoptotic HCC38 cells for HLA subpopulations after 72 h treatment with 15.6 nM doxorubicin. (**h**) Fraction of remaining cell population consisting of HLA$^{lo}$ and HLA$^{hi}$ HCC38 cells after 72 h doxorubicin treatment. For (**c**-**h**), data represents $n$ = 4 biologically independent samples. Error bars are mean ± s.e.m., and $P$ values calculated using a two-sided $t$-test.

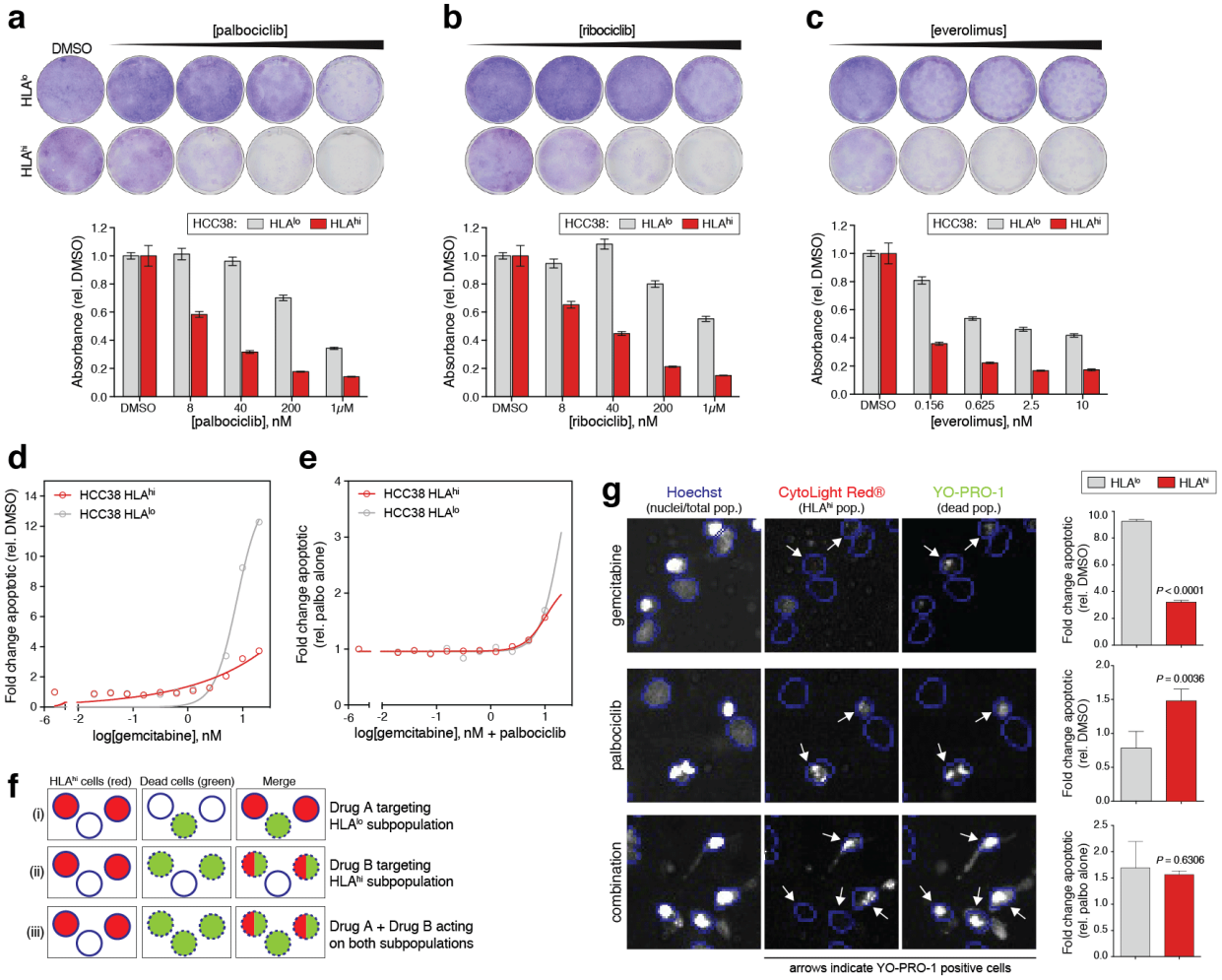
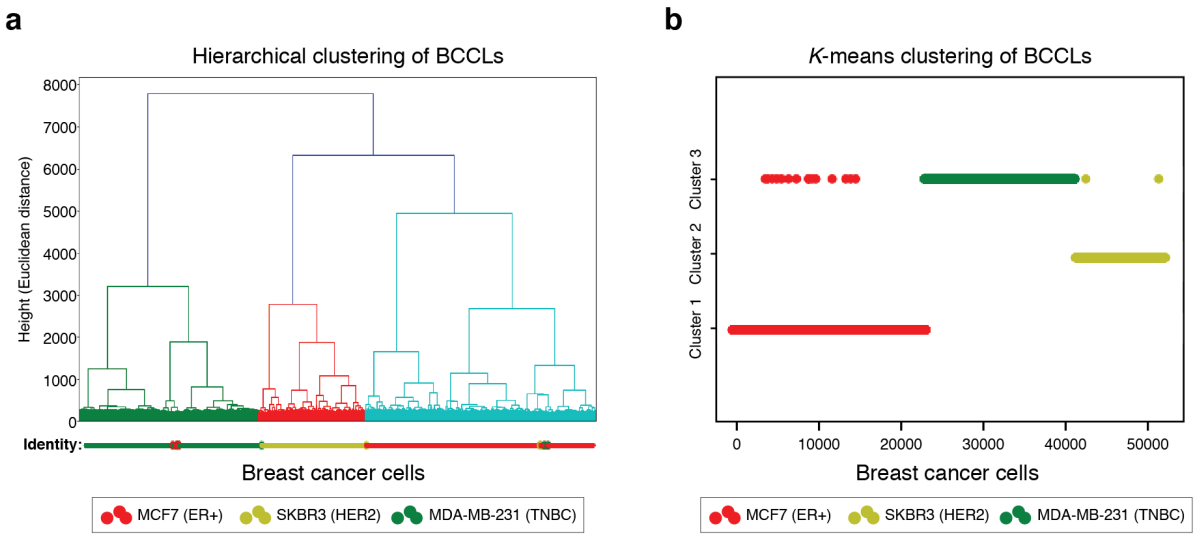
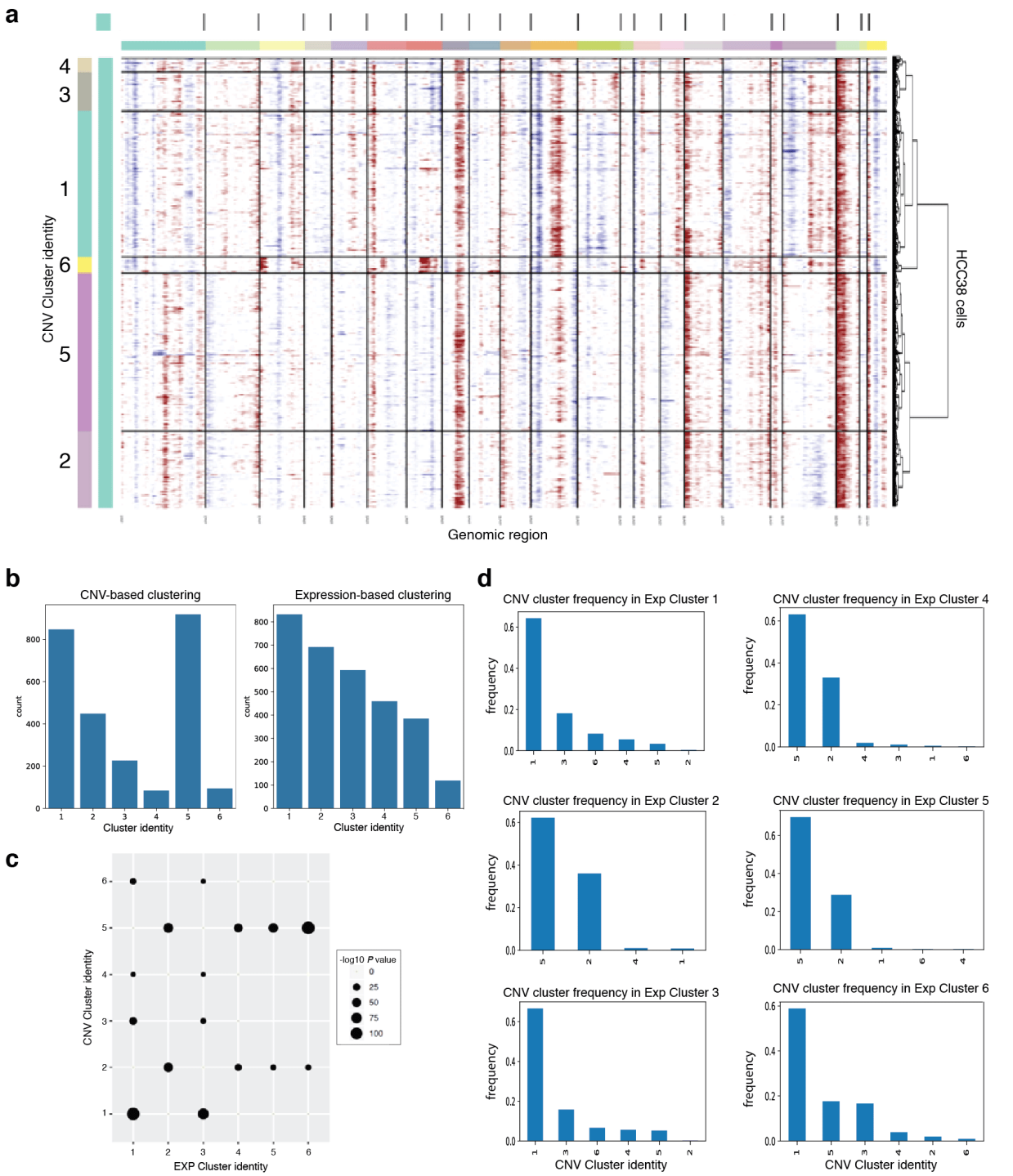



**Figure 3.7 Targeting a subpopulation of cells in an ISG-state with CDK4 or mTOR inhibitors increases efficacy of chemotherapies on the full tumor landscape.** Crystal violet staining of HLA$^{lo}$ and HLA$^{hi}$ HCC38 cells after 14 d treatment with increasing concentrations of (**a**) palbociclib (CDK4i), (**b**) ribociclib (CDK4/6i), and (**c**) everolimus (mTORi). Representative images of $n$ = 2 independent experiments with similar results, and quantification of dye shown below. Fold change apoptotic cells for HCC38 HLA subpopulations after 72 h treatment with (**d**) single-agent gemcitabine or (**e**) in combination with 0.5 μM palbociclib. Combinations were normalized to palbociclib alone. (**f**) Schematic showing drugs acting on distinct subpopulations of cells. (**g**) Representative images showing 10 nM gemcitabine inducing cell death in HCC38 HLA$^{lo}$ cells (top), 0.5 μM palbociclib inducing cell death in HCC38 HLA$^{hi}$ cells (middle), and the combination of the two agents resulting in cell death in both subpopulations (bottom). Arrows indicate YO-PRO-1-positive (dead) cells. Quantification shown to the right. In all graphs, data represents $n$ = 4 independent samples, error bars are mean ± s.d., and $P$ values calculated using a two-sided $t$-test unless stated otherwise.

## 3.9 Supplementary Figures

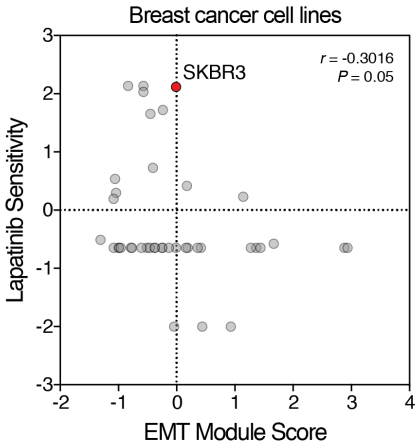



**Supplementary Figure 3.1 Methods for measuring cell-cell similarity and clustering.** (**a**) Hierarchical clustering of MCF7 (ER+), SKBR3 (HER2) and MDA-MB-231 (TNBC) cells from single-cell RNAseq profiling data is shown. Dendogram was constructed using Ward Linkage with squared Euclidean distance as the unit of measure between points. Individual cells are annotated based on parent cell line identity shown below. (**b**) *K*-means clustering result for individual breast cancer cells. Individual cells fall into three dominant clusters and are annotated based on parent cell line identity as in panel (a). BCCLs, breast cancer cell lines.

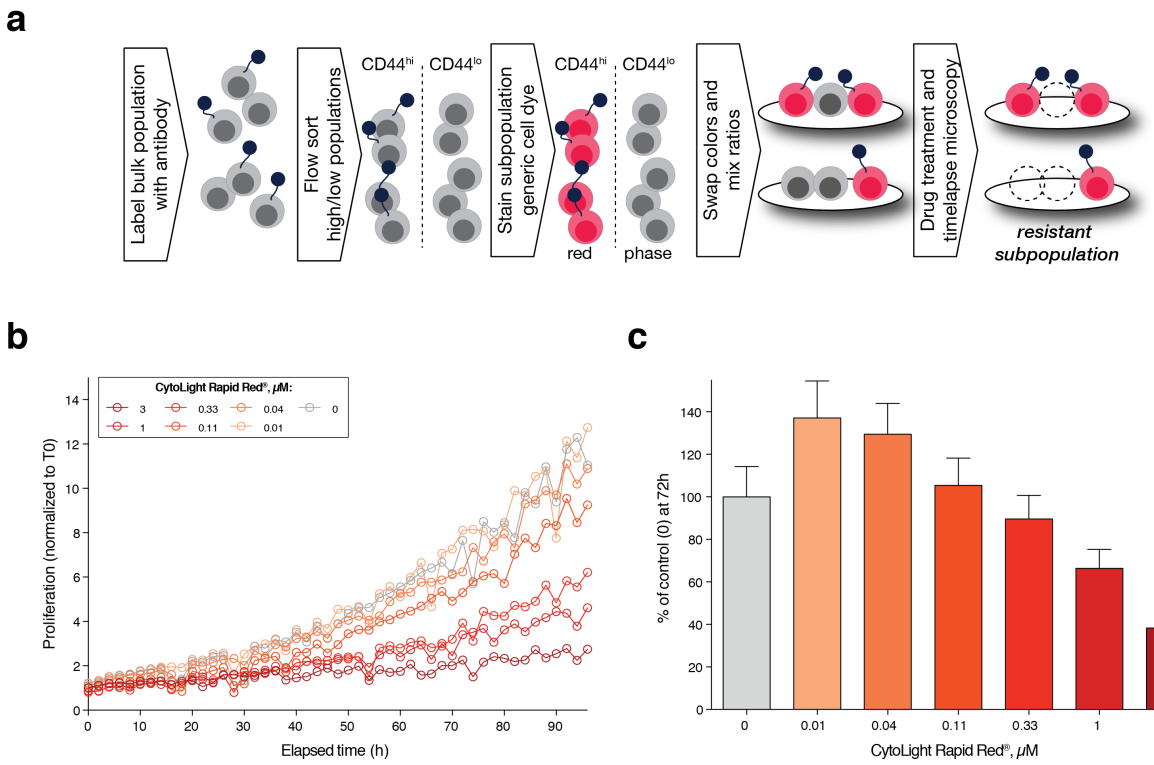



**Supplementary Figure 3.2 Interpreting expression-based subpopulations of HCC38 cells using CNV estimations.** (**a**) Chromosomal landscape of inferred large-scale CNVs for HCC38 cells as an output from running InferCNV[31] Individual HCC38 cells (y axis) and chromosomal regions (x axis) are shown with amplifications (red) or deletions (blue) inferred by averaging expression over 100-gene stretches on the respective chromosomes. (**b**) Comparison of number of cells

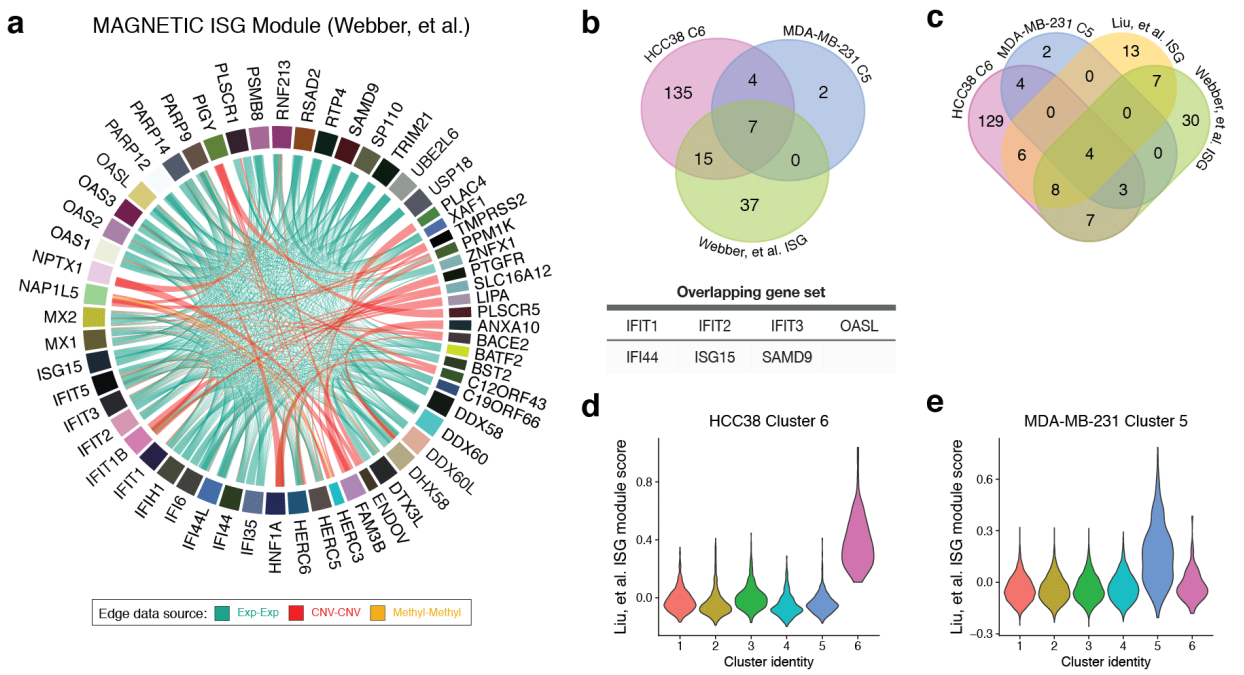
within each subgroup identified using CNV-based and expression-based clustering approaches. (**c**) Enrichment plot showing expression-based clusters belong to multiple genetic subclones. (**d**) Frequency of CNV clusters present in each expression-based cluster of cells.



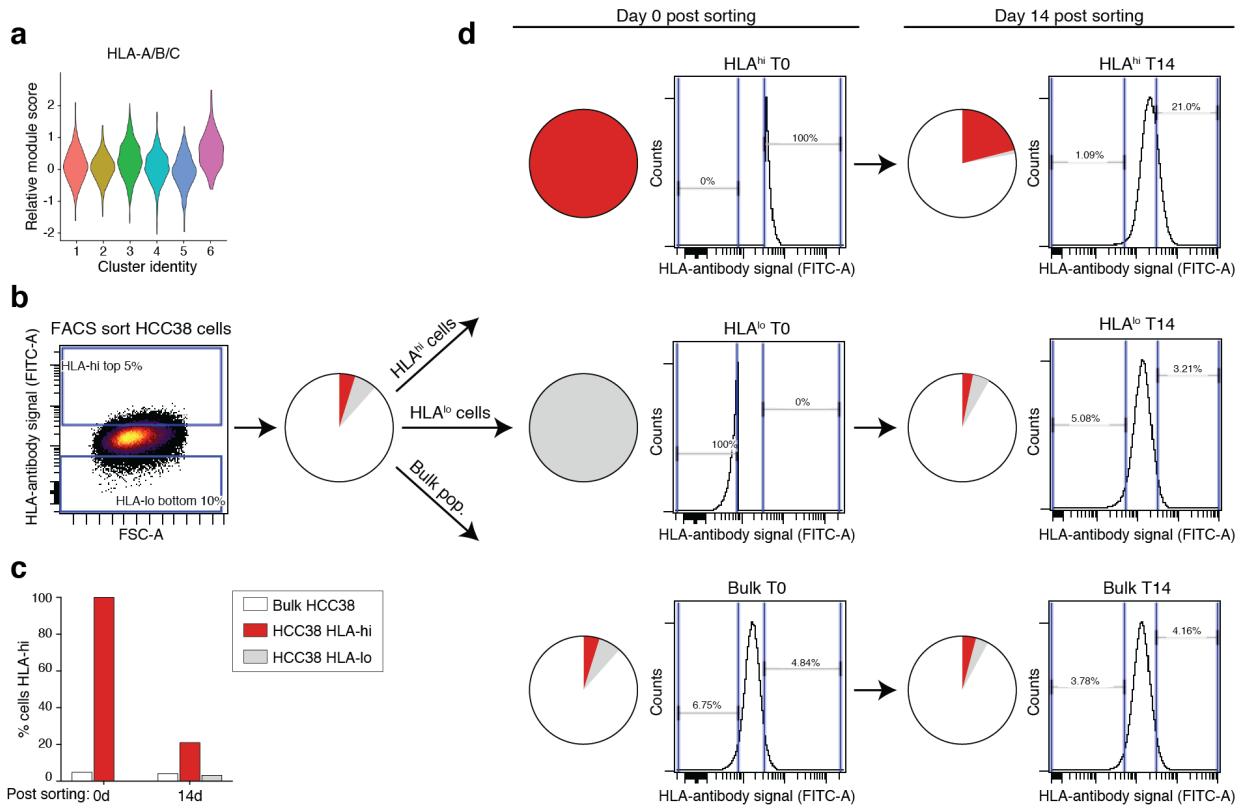

**Supplementary Figure 3.3 EMT module status correlates with lapatinib sensitivity.** A panel of 82 breast cancer cell lines were scored based on EMT module status. Scatter plot shows EMT-module scores negatively associate with lapatinib sensitivity based on $-\log_{10}$ GI50 values from published datasets (Heiser, et al., Daemen, et al.)[40,41]. The HER2-amplifed cell line SKBR3 is highlighted in red. *P* value based on Pearson correlation.



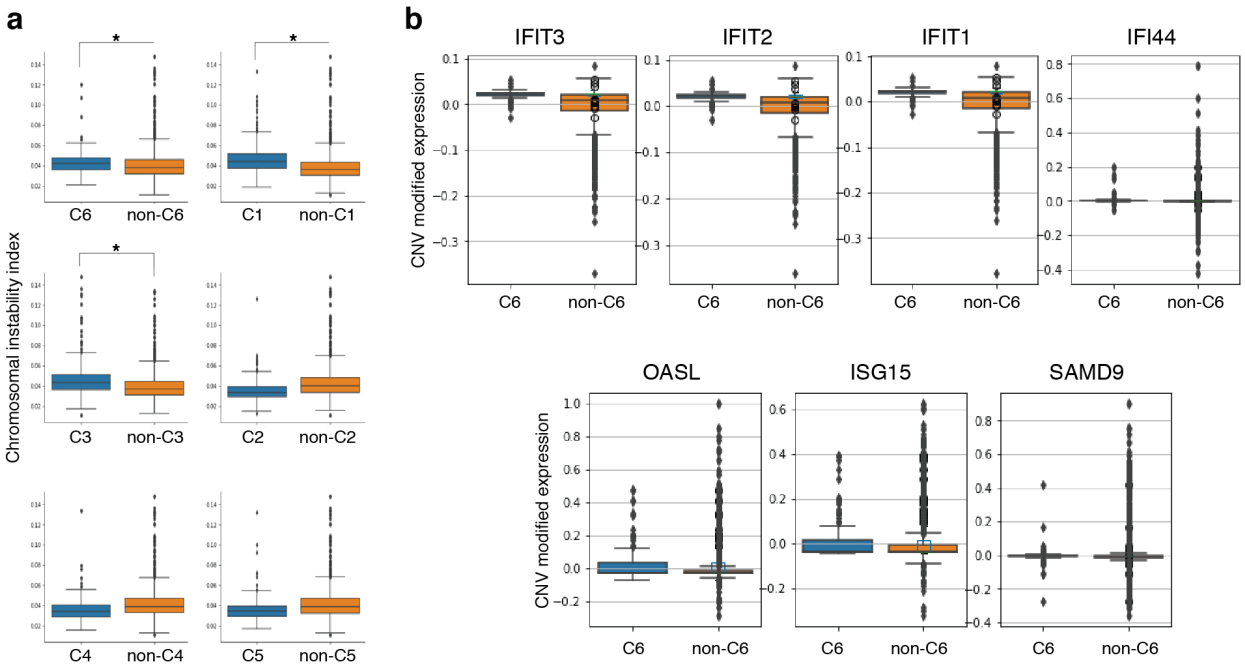

**Supplementary Figure 3.4 Single-live-cell tracking optimization for SKBR3 cells.** (**a**) Single-live-cell tracking workflow scheme. Bulk SKBR3 cells are labeled with CD44 antibody. CD44$^{hi}$ and CD44$^{lo}$ subpopulations are isolated via FACS. CD44$^{hi}$ cells are stained red with 0.11 μM of CytoLight Rapid Red® reagent to track subpopulation over time. CD44$^{lo}$ cells are left unstained. CD44$^{hi}$-red and CD44$^{lo}$-unlabeled subpopulations are mixed back into single bulk model, treated with drug for 72 h, and imaged using time-lapse microscopy. (**b**) Proliferation of SKBR3 cells over time following increasing concentrations of CytoLight Rapid Red® reagent. (**c**) Relative fraction of SKBR3 cells expressed as a percentage of control (DMSO) after 72 h staining with increasing concentrations of CytoLight Rapid Red®. For (**b,c**) data represents $n$ = 4 biologically independent samples and error bars are mean ± s.d.



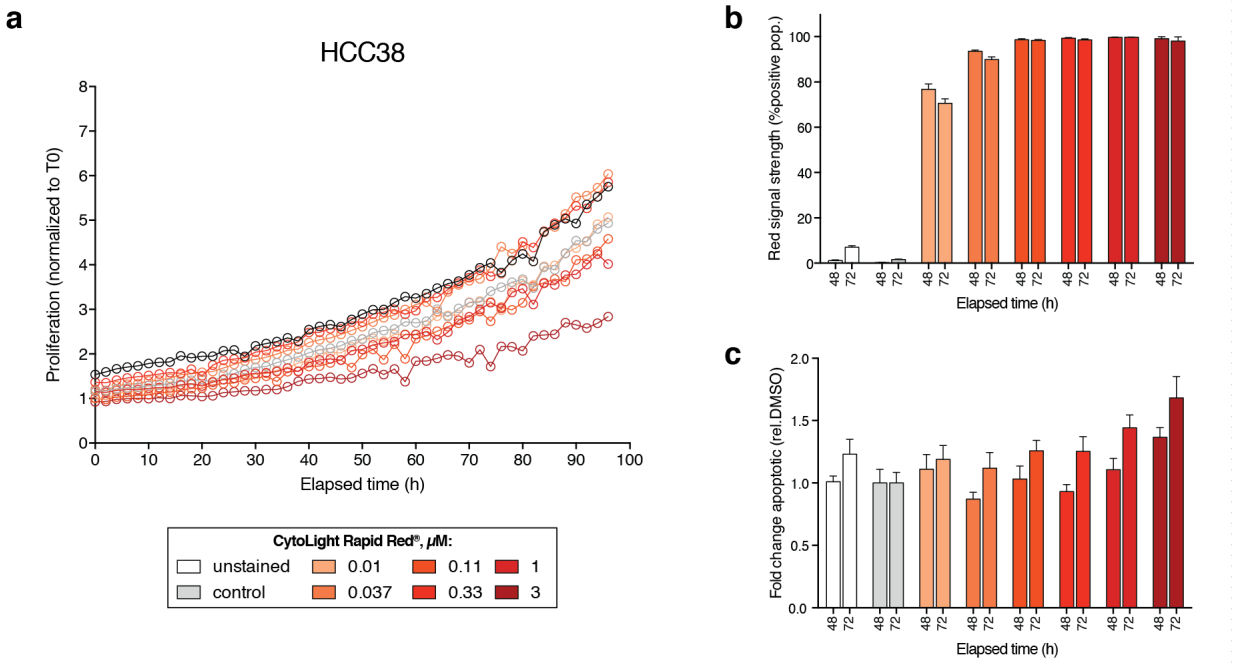

**Supplementary Figure 3.5 ISG signature overlap in TNBC cells.** (**a**) Circos plot showing genes comprising the ISG module derived using the MAGNETIC algorithm (Webber, et al.)[37]. Linkages are colored based on edge source data type. (**b**) Overlap of genes involved in the ISG module identified in Webber, et al. in comparison to gene lists encompassing significant differentially expressed genes from HCC38 Cluster 6 and MDA-MB-231 Cluster 5 cells (relative to all other clusters). Genes overlapping in all three gene sets are shown below. (**c**) Overlap of genes identified in published ISG signatures (Webber, et al., Liu, et al.)[37,42] in comparison to differentially expressed gene lists from HCC38 Cluster 6 and MDA-MB-231 Cluster 5 cells. Relative ISG module scores using the Liu, et al. signature shown for individual (**d**) HCC38 and (**e**) MDA-MB-231 cells.



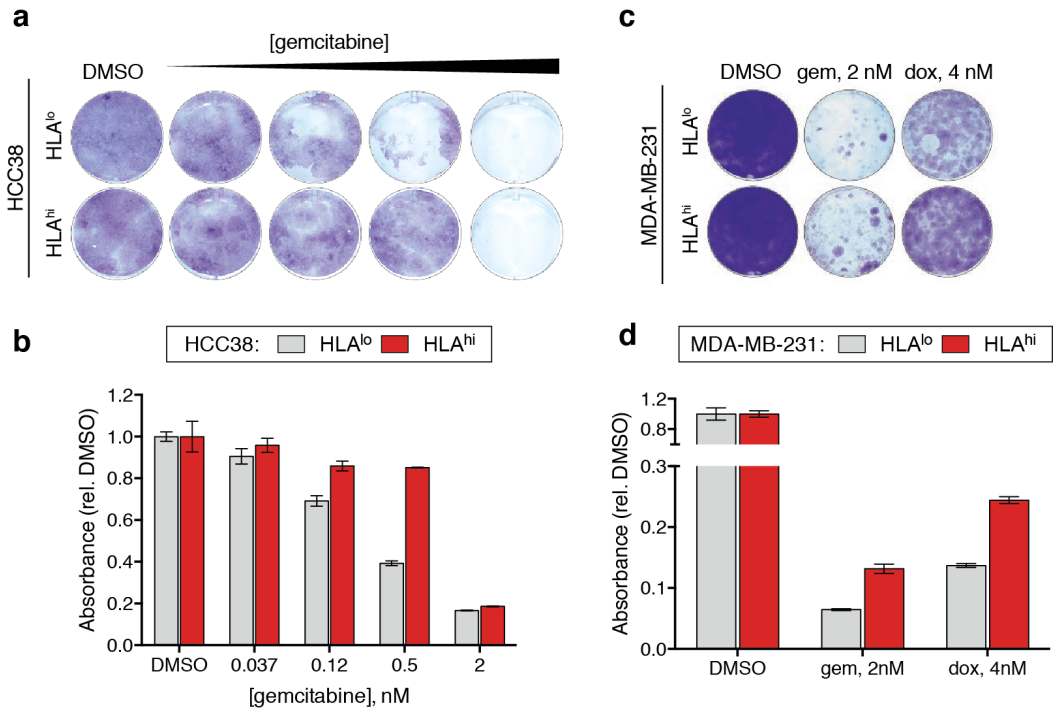

**Supplementary Figure 3.6 HLA kinetics in HCC38 cells.** (**a**) *HLA-A*, *HLA-B*, and *HLA-C* were compiled into a single gene module (HLA-A/B/C) and HCC38 cells were scored based on their expression of module genes. Violin plot of the relative gene module score for individual cells within each subcluster is shown. Top significantly differentially expressed gene lists from. (**b**) Isolation via FACS of HCC38 HLA-hi (top 5%) and HLA-lo (bottom 10%) populations. (**c**) Percentage of cells in each subpopulation identified as HLA-hi via FACS analysis immediately after sorting (0 d) and 14 d post initial sort. (**d**). HLA kinetics in various HCC38 subpopulations is shown over 14 d. Pie charts depict relative proportions of HLA-hi (red) and HLA-lo (grey) subpopulations for an unsorted sample of HCC38 cells (left, initial distribution) and for sorted subpopulations (top, HLA-high; middle, HLA-lo; bottom, control bulk population) analyzed immediately after sorting (0 d) and 14 d post initial sort. Data represents *n* = 2 independent experiments.




**Supplementary Figure 3.7 HCC38 ISG cells have heightened genomic instability as determined by CNV analysis.** (**a**) Chromosome instability shown for HCC38 cells in each expression-based subpopulation compared to remaining clusters. *P* value calculated using a two-sided *t*-test. (**b**) Interferon-stimulated genes have significantly higher CNV levels in HCC38 cells in the ISG subpopulation (Cluster 6) compared to non-Cluster 6 cells. In all graphs, Box-and-whisker plots show median (centerline), upper/lower quartiles (box limits), and whiskers spanning the interquartile range from 25-75 percentiles. CNV, copy number variation.




**Supplementary Figure 3.8 Single-live-cell tracking optimization for HCC38 cells.** (**a**) Proliferation of HCC38 cells over time following increasing concentrations of CytoLight Rapid Red® reagent. (**b**) Red signal strength shown as percentage of positive cells with detectable signal at 48 and 72 h after staining with the indicated concentrations of CytoLight Rapid Red® reagent. (**c**) Fold change apoptotic cells following 48 and 72 h staining with increasing concentrations of CytoLight Rapid Red® reagent. All data represents $n$ = 4 biologically independent samples and error bars are mean ± s.d.




**Supplementary Figure 3.9 Subpopulations of TNBC cells in an ISG-state have increased long-term survival in response to chemotherapies.** (**a**) Crystal violet staining of HLA$^{hi}$ (top 5%) and HLA$^{lo}$ (remaining) HCC38 cells after 14 d treatment with increasing concentrations of gemcitabine. (**b**) Quantification of dye from images shown in panel (**a**). (**c**) Crystal violet staining of MDA-MB-231 HLA$^{hi}$ (top 5%) and HLA$^{lo}$ (remaining) subpopulations following 14 d treatment with 2 nM gemcitabine or 4 nM doxorubicin. (**d**) Quantification of dye from images shown in panel (**c**). Images are representative of $n$ = 2 independent experiments with similar results. In all graphs, data represents $n$ = 4 independent samples.

## 3.10 References


1. Cortazar, P. *et al.* Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet Lond. Engl.* **384**, 164–172 (2014).

2. von Minckwitz, G. *et al.* Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **30**, 1796–1804 (2012).

3. Symmans, W. F. *et al.* Long-Term Prognostic Risk After Neoadjuvant Chemotherapy Associated With Residual Cancer Burden and Breast Cancer Subtype. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **35**, 1049–1060 (2017).

4. Foulkes, W. D., Smith, I. E. & Reis-Filho, J. S. Triple-negative breast cancer. *N. Engl. J. Med.* **363**, 1938–1948 (2010).

5. Liedtke, C. *et al.* Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **26**, 1275–1281 (2008).

6. Untch, M. *et al.* Neoadjuvant treatment with trastuzumab in HER2-positive breast cancer: results from the GeparQuattro study. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **28**, 2024–2031 (2010).

7. Pierga, J.-Y. *et al.* A multicenter randomized phase II study of sequential epirubicin/cyclophosphamide followed by docetaxel with or without celecoxib or trastuzumab according to HER2 status, as primary chemotherapy for localized invasive breast cancer patients. *Breast Cancer Res. Treat.* **122**, 429–437 (2010).

8. Bhang, H. C. *et al.* Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. *Nat. Med.* **21**, 440–448 (2015).

9. Kim, C. *et al.* Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing. *Cell* **173**, 879-893.e13 (2018).

10. Sharma, A. *et al.* Longitudinal single-cell RNA sequencing of patient-derived primary cells reveals drug-induced infidelity in stem cell hierarchy. *Nat. Commun.* **9**, 1–17 (2018).

11. Yates, L. R. *et al.* Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat. Med.* **21**, 751–759 (2015).

12. Balko, J. M. *et al.* Molecular profiling of the residual disease of triple-negative breast cancers after neoadjuvant chemotherapy identifies actionable therapeutic targets. *Cancer Discov.* **4**, 232–245 (2014).

13. Almendro, V. *et al.* Inference of tumor evolution during chemotherapy by computational modeling and in situ analysis of genetic and phenotypic cellular diversity. *Cell Rep.* **6**, 514–527 (2014).

14. Echeverria, G. V. *et al.* Resistance to neoadjuvant chemotherapy in triple-negative breast cancer mediated by a reversible drug-tolerant state. *Sci. Transl. Med.* **11**, (2019).

15. Shaffer, S. M. *et al.* Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature* **546**, 431–435 (2017).

16. Risom, T. *et al.* Differentiation-state plasticity is a targetable resistance mechanism in basal-like breast cancer. *Nat. Commun.* **9**, 1–17 (2018).

17. Creighton, C. J. *et al.* Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13820–13825 (2009).

18. Lehmann, B. D. *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* **121**, 2750–2767 (2011).

19. Gupta, P. B. *et al.* Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146**, 633–644 (2011).

20. Ayers, M. *et al.* Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **22**, 2284–2293 (2004).

21. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).

22. Sørlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10869–10874 (2001).

23. Rivenbark, A. G., O'Connor, S. M. & Coleman, W. B. Molecular and cellular heterogeneity in breast cancer: challenges for personalized medicine. *Am. J. Pathol.* **183**, 1113–1124 (2013).

24. Chung, W. *et al.* Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat. Commun.* **8**, 1–12 (2017).

25. Savage, P. *et al.* A Targetable EGFR-Dependent Tumor-Initiating Program in Breast Cancer. *Cell Rep.* **21**, 1140–1149 (2017).

26. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *ArXiv180203426 Cs Stat* (2018).

27. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).

28. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).

29. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).

30. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).

31. Tickle, T., Tirosh, I., Georgescu, C., Brown, M. & Haas, B. *InferCNV of the Trinity CTAT Project*. (Broad Institute of MIT and Harvard, 2019).

32. Huper, G. & Marks, J. R. Isogenic normal basal and luminal mammary epithelial isolated by a novel method show a differential response to ionizing radiation. *Cancer Res.* **67**, 2990–3001 (2007).

33. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).

34. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).

35. Simpson, E. H. The Interpretation of Interaction in Contingency Tables. *J. R. Stat. Soc. Ser. B Methodol.* **13**, 238–241 (1951).

36. Parker, J. S. *et al.* Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).

37. Webber, J. T., Kaushik, S. & Bandyopadhyay, S. Integration of Tumor Genomic Data with Cell Lines Using Multi-dimensional Network Modules Improves Cancer Pharmacogenomics. *Cell Syst.* **7**, 526-536.e6 (2018).

38. Creedon, H. *et al.* Identification of novel pathways linking epithelial-to-mesenchymal transition with resistance to HER2-targeted therapy. *Oncotarget* **7**, 11539–11552 (2016).

39. Lesniak, D. *et al.* Spontaneous epithelial-mesenchymal transition and resistance to HER-2-targeted therapies in HER-2-positive luminal breast cancer. *PloS One* **8**, e71987 (2013).

40. Heiser, L. M. *et al.* Subtype and pathway specific responses to anticancer compounds in breast cancer. *Proc. Natl. Acad. Sci.* **109**, 2724–2729 (2012).

41. Daemen, A. *et al.* Modeling precision treatment of breast cancer. *Genome Biol.* **14**, R110 (2013).

42. Liu, H. *et al.* Tumor-derived IFN triggers chronic pathway agonism and sensitivity to ADAR loss. *Nat. Med.* **25**, 95–102 (2019).

43. Wolf, D. M., Lenburg, M. E., Yau, C., Boudreau, A. & Veer, L. J. van 't. Gene Co-Expression Modules as Clinically Relevant Hallmarks of Breast Cancer Diversity. *PLOS ONE* **9**, e88309 (2014).

44. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).

45. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).

46. Borden, E. C. Interferons α and β in cancer: therapeutic opportunities from new insights. *Nat. Rev. Drug Discov.* **18**, 219–234 (2019).

47. Mackenzie, K. J. *et al.* cGAS surveillance of micronuclei links genome instability to innate immunity. *Nature* **548**, 461–465 (2017).

48. Boelens, M. C. *et al.* Exosome Transfer from Stromal to Breast Cancer Cells Regulates Therapy Resistance Pathways. *Cell* **159**, 499–513 (2014).

49. Härtlova, A. *et al.* DNA Damage Primes the Type I Interferon System via the Cytosolic DNA Sensor STING to Promote Anti-Microbial Innate Immunity. *Immunity* **42**, 332–343 (2015).

50. Zevini, A., Olagnier, D. & Hiscott, J. Cross-Talk between the Cytoplasmic RIG-I and STING Sensing Pathways. *Trends Immunol.* **38**, 194–205 (2017).

51. Duijf, P. H. G. *et al.* Mechanisms of Genomic Instability in Breast Cancer. *Trends Mol. Med.* **25**, 595–611 (2019).

52. Lee, H. J. *et al.* Differential expression of major histocompatibility complex class I in subtypes of breast cancer is associated with estrogen receptor and interferon signaling. *Oncotarget* **7**, 30119–30132 (2016).

53. Sporikova, Z., Koudelakova, V., Trojanec, R. & Hajduch, M. Genetic Markers in Triple-Negative Breast Cancer. *Clin. Breast Cancer* **18**, e841–e850 (2018).

54. Shah, S. P. *et al.* The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486**, 395–399 (2012).

55. Gaston, J. *et al.* Intracellular STING inactivation sensitizes breast cancer cells to genotoxic agents. *Oncotarget* **7**, 77205–77224 (2016).

56. Weichselbaum, R. R. *et al.* An interferon-related gene signature for DNA damage resistance is a predictive marker for chemotherapy and radiation for breast cancer. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 18490–18495 (2008).

57. Roden, D. L. *et al.* Single cell transcriptomics reveals molecular subtype and functional heterogeneity in models of breast cancer. *bioRxiv* 282079 (2018) doi:10.1101/282079.

58. Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**, 515–527 (2006).

59. DeRose, Y. S. *et al.* Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. *Nat. Med.* **17**, 1514–1520 (2011).

60. Camarda, R. *et al.* Inhibition of fatty acid oxidation as a therapy for MYC-overexpressing triple-negative breast cancer. *Nat. Med.* **22**, 427–432 (2016).

61. Lawson, D. A. *et al.* Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. *Nature* **526**, 131–135 (2015).

62. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).

63. Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (2014).

# Chapter 4.

# Conclusions and future directions

Here we developed two approaches towards designing novel combination strategies for overcoming drug resistance in breast cancer. First, using an unbiased chemoproteomics approach, we profiled kinome dynamics across breast cancer cells in response to various targeted therapies and identified signaling changes that correlate with drug sensitivity. This signaling map identified survival factors whose presence limits the efficacy of targeted therapies and revealed AURKA as a new co-targeting opportunity to enhance the therapeutic efficacy of PI3K-pathway inhibitors in breast cancer. The systematic measurement of kinome dynamics across a diverse set of cell lines allowed us to map molecular changes associated with resistance to a variety of inhibitors, which is unique from previous approaches limited to a single drug or cell line. We found a number of cases in which failure to inhibit a particular kinase was associated with drug resistance. As our proteomic screen included multiple drugs that impinge on distinct oncogenic pathways, we found it surprising that a set of common survival factors were identified. Future work may determine if other candidates we identified also act as survival factors and how they might do so. We also showed that MYC-driven AURKA activation results in maintenance of the PI3K-pathway despite PI3K inhibitor treatment, resulting in drug resistance. Future work may gauge the relative importance of AURKA in respect to other outputs of MYC in driving resistance to PI3K inhibitors. Our findings reveal that the combination of Aurora kinase inhibitors and PI3K-pathway inhibitors is synergistic and could be a promising clinical strategy

to enhance the treatment response in breast cancer. As single-agent responses to both PI3K-pathway and Aurora kinase inhibitors have been modest, these findings may unlock the full potential of these agents in realizing a clinical benefit.

Second, we used single-cell transcriptomics data and pharmacogenomic modeling as a way to inform upfront drug combinations based on systematic analysis of tumor subpopulation architectures. Our data identified a recurrent subpopulation of cells in an ISG-state that was specific to TNBC cells and displayed heightened STING-pathway activation and genomic instability. By applying our $MOD_{Rx}$ approach we determined these cells are associated with chemoresistance and poor outcome in TNBC patients, but can be targeted using CDK4/6 or mTOR inhibitors. Our study potentially explains the cellular source of heightened ISG expression found in residual TNBC tumors after chemotherapy and calls for future work exploring the underlying mechanisms and clinical utility of targeting pre-existing chemoresistant ISG-subpopulations in TNBC. Using *in silico* and experimental approaches, our study provides an effective new framework to discover drug combinations capable of counteracting intrinsic cell variability by predicting drug responses of single cells within tumor cell subpopulations and systematically links transcriptional heterogeneity with drug actionability to optimize therapy combinations. Ongoing efforts are currently underway to translate this approach at scale by developing a machine-learning model for generating a drug-by-cell matrix to provide a systematic method for predicting single-cell responses and inform candidate combination strategies.

which we validated as an effective strategy for targeting this chemoresistant subpopulation in TNBC. Guided by our *in silico* modeling and single-cell predictions, we show the addition of palbociclib to gemcitabine or doxorubicin treatment is an effective strategy for increasing therapeutic efficacy on TNBC tumor landscapes in their entirety.

**Publishing Agreement**

*It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.*

***Please sign the following statement:***

*I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.*

DocuSigned by:
Hayley Donnella
C93B3E70BAB844C...
Author Signature

12/10/2019
Date